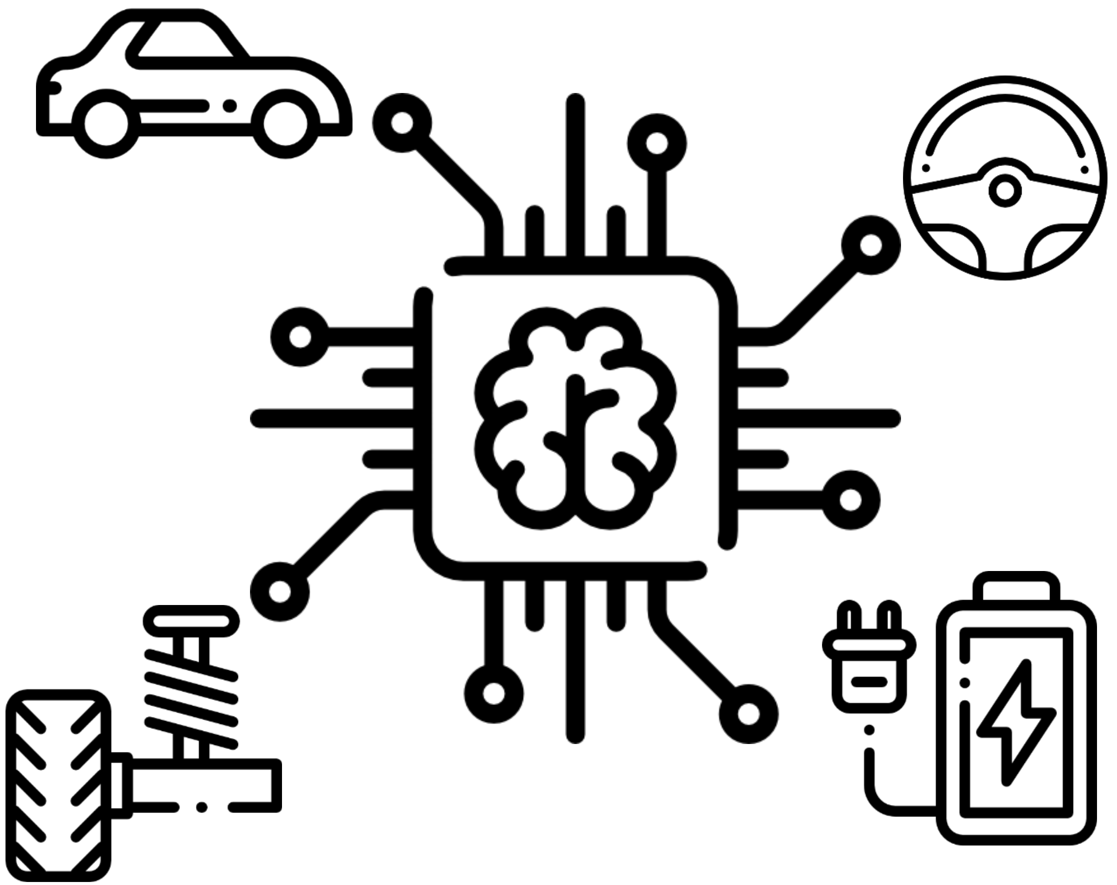
Lehrstuhl für Fahrzeugtechnik
Fakultät für Maschinenwesen
Technische Universität München

Artificial Intelligence in Automotive Technology

Johannes Betz / Prof. Dr.-Ing. Markus Lienkamp / Prof. Dr.-Ing. Boris Lohmann



Lecture Overview

1 Introduction: Artificial Intelligence 18.10.2018 – Johannes Betz	**6 Pathfinding: From British Museum to A*** 29.11.2018 – Lennart Adenaw	**11 Reinforcement Learning** 17.01.2019 – Christian Dengler
Practice 1 18.10.2018 – Johannes Betz	**Practice 6** 29.11.2018 – Lennart Adenaw	**Practice 11** 17.01.2019 – Christian Dengler
2 Perception 25.10.2018 – Johannes Betz	**7 Introduction: Artificial Neural Networks** 06.12.2018 – Lennart Adenaw	**12 AI-Development** 24.01.2019 – Johannes Betz
Practice 2 25.10.2018 – Johannes Betz	**Practice 7** 06.12.2018 – Lennart Adenaw	**Practice 12** 24.01.2019 – Johannes Betz
3 Supervised Learning: Regression 08.11.2018 – Alexander Wischnewski	**8 Deep Neural Networks** 13.12.2018 – Jean-Michael Georg	**13 Free Discussion** 31.01.2019 – Betz/Adenaw
Practice 3 08.11.2018 – Alexander Wischnewski	**Practice 8** 13.12.2018 – Jean-Michael Georg	
4 Supervised Learning: Classification 15.11.2018 – Jan Cedric Mertens	**9 Convolutional Neural Networks** 20.12.2018 – Jean-Michael Georg	
Practice 4 15.11.2018 – Jan Cedric Mertens	**Practice 9** 20.12.2018 – Jean-Michael Georg	
5 Unsupervised Learning: Clustering 22.11.2018 – Jan Cedric Mertens	**10 Recurrent Neural Networks** 10.01.2019 – Christian Dengler	
Practice 5 22.11.2018 – Jan Cedric Mertens	**Practice 10** 10.01.2019 – Christian Dengler	

Feedback from last week

- Repeating the Quiz – Yes, its possible now!

Objectives for Lecture 5: Clustering

After the lecture you are able to…

… understand the concept of clustering and its association to pattern recognition.

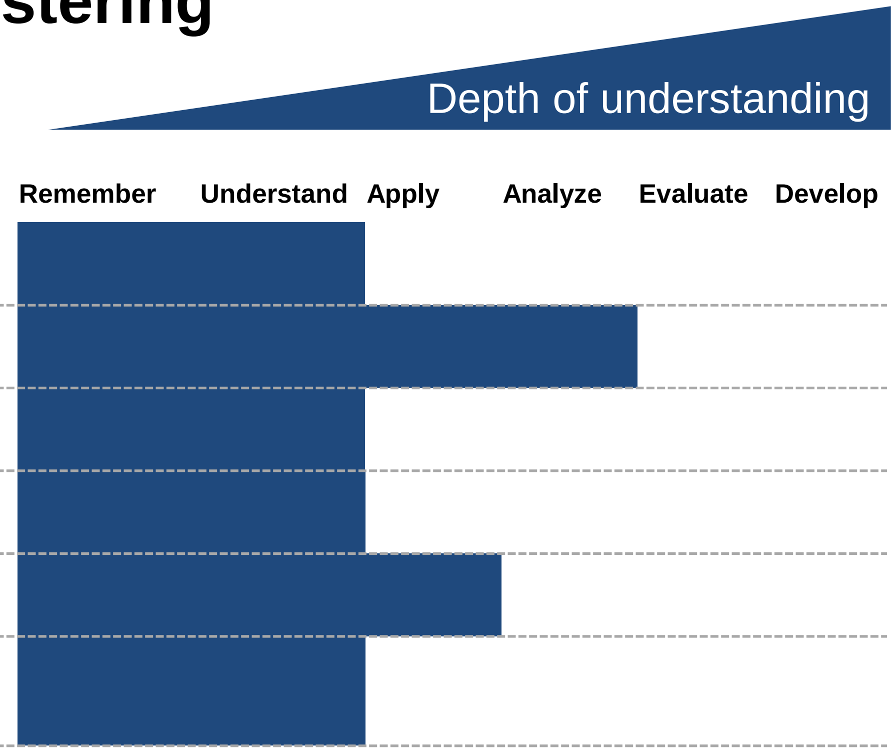
… analyze the quality of given clusters regarding to different criteria.

… understand the workflow of unsupervised learning.

… understand the concepts of different clustering methods together with their pro and cons.

… implement, train and use a clustering method with python libraries.

… identify if a problem belongs to regression, classification or clustering.



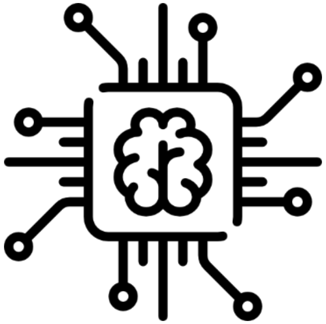
Supervised Learning: Classification

Johannes Betz / Prof. Dr. Markus Lienkamp / Prof. Dr. Boris Lohmann

(Jan Cedric Mertens, M.Sc.)

Agenda

1. Chapter: Introduction
    - **1.1 Overview**
    - 1.2 Training and Validation
2. Chapter: Methods
    - 2.1 Hierarchical Clustering
    - 2.2 k-means
    - 2.3 DBSCAN
3. Chapter: Application
4. Chapter: Summary



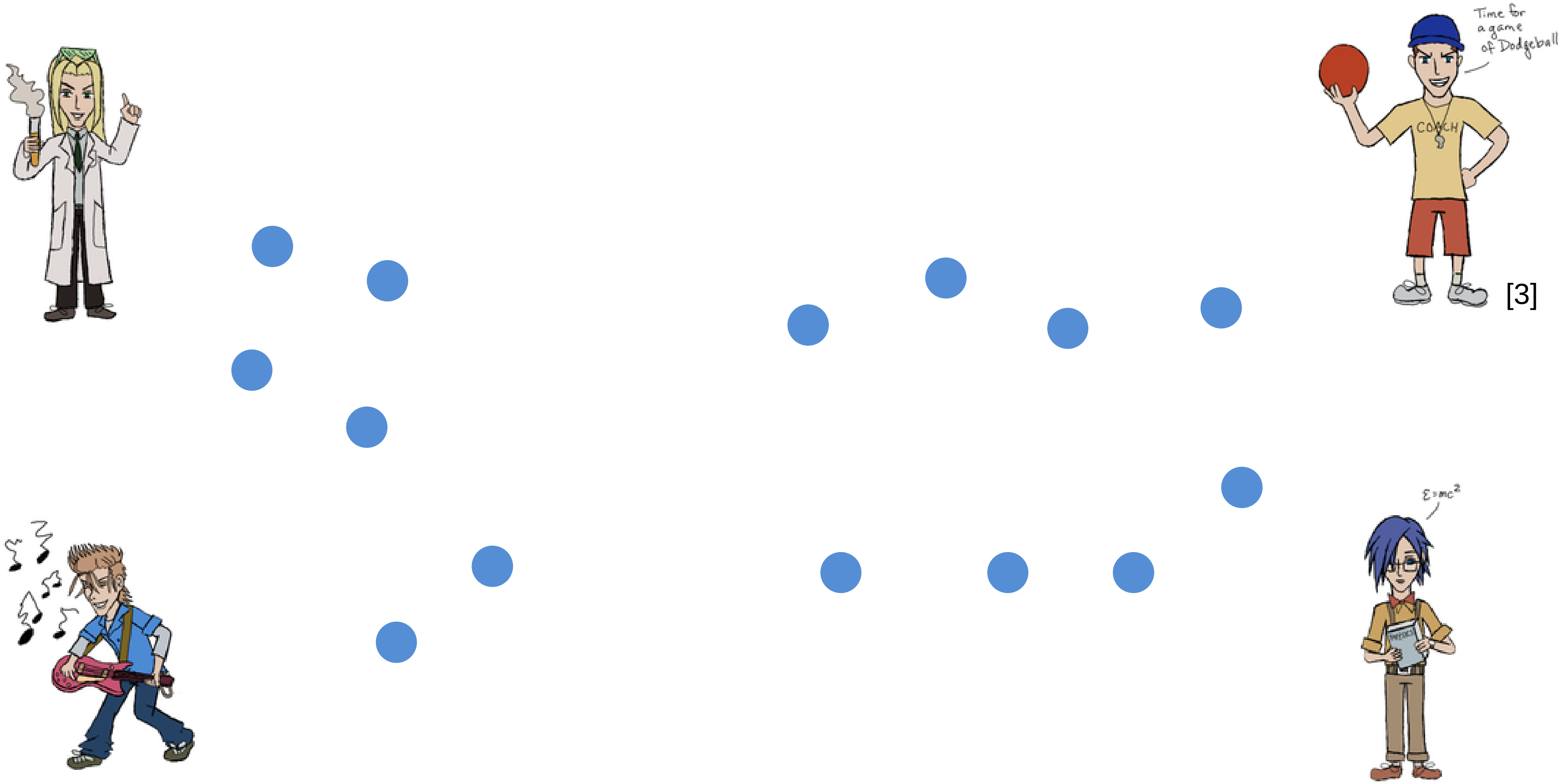
Clustering

"Grouping of similar things that are close together, sometimes surrounding something" [2]

[1]

Clustering

"Grouping of similar things that are close together, sometimes surrounding something" [2]



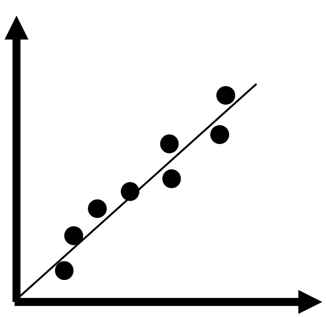
[3]

Method Overview

Pattern Recognition

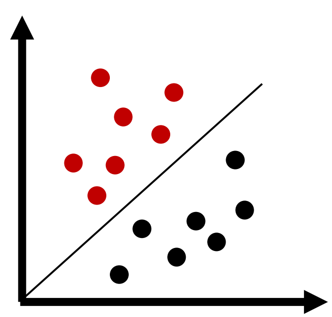
Regression

- Predict **continuous** valued output
- Supervised



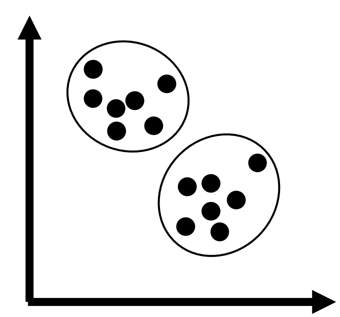
Classification

- Predict **discrete** valued output
- Supervised

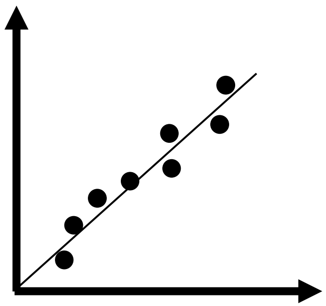
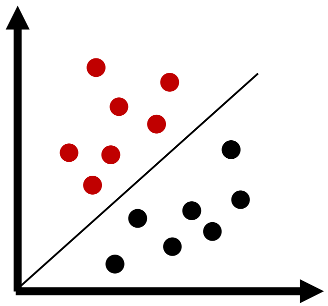
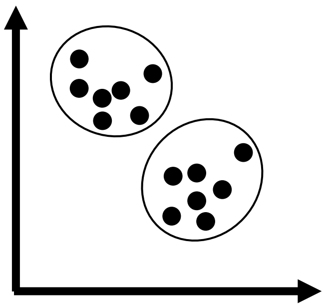


Clustering

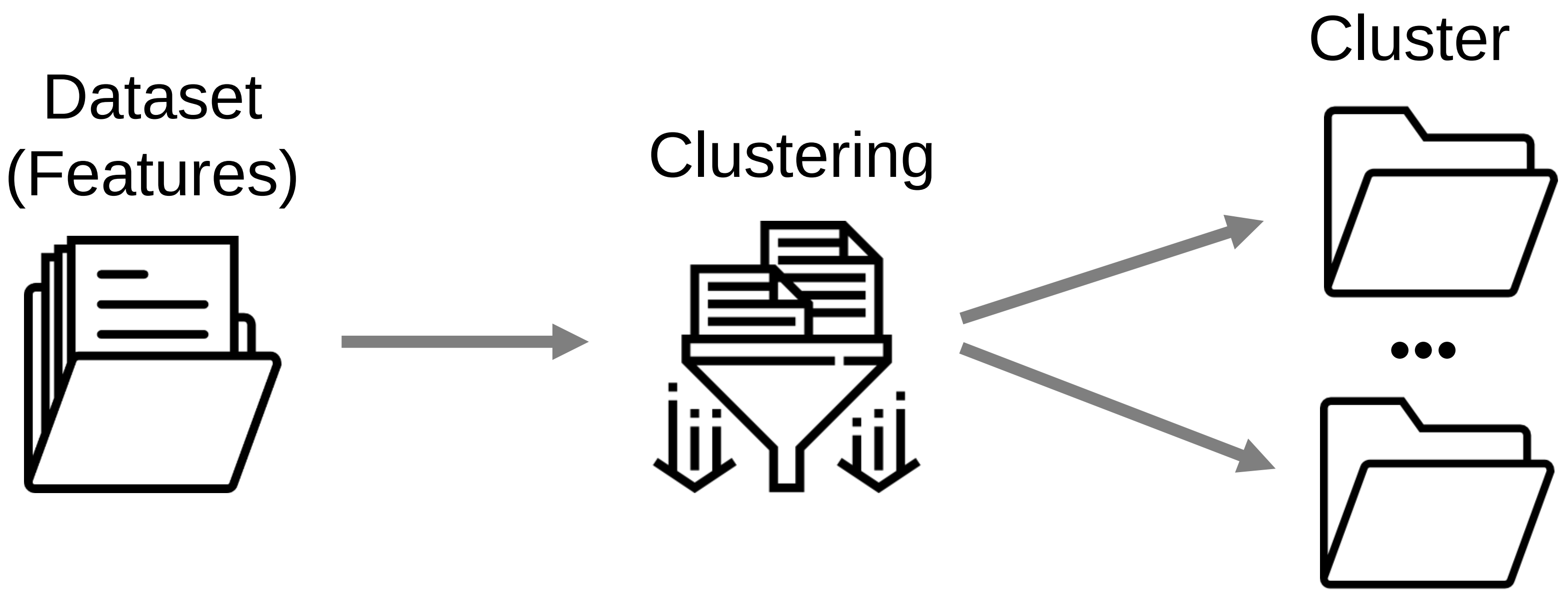
- Predict discrete valued output
- **Unsupervised**



Method Overview

Regression	Classification	Clustering
		
▪ House pricing ▪ Number of sales ▪ Persons weight	▪ Object detection ▪ Spam detection ▪ Cancer detection	▪ Genome patterns ▪ Google news ▪ Pointcloud (Lidar) processing

General Approach



News (Keywords, …)	Similarities?	Cluster
Genomes (Size, …)	Similarities?	Cluster
Points (Position, …)	Similarities?	Cluster

Clustering - Example

Cluster Points → Classifiy subset of Cluster → Assign class to cluster

[4]

Clustering vs. Segmentation

- Both terms are interchangable
- Statistical background: Clustering
- Business background: Segment
- Clustering produces segments and vice versa

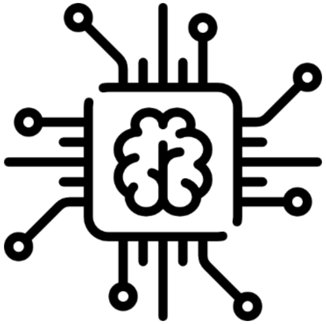
Supervised Learning: Classification

Johannes Betz / Prof. Dr. Markus Lienkamp / Prof. Dr. Boris Lohmann

(Jan Cedric Mertens, M.Sc.)

Agenda

1. Chapter: Introduction
    - 1.1 Overview
    - **1.2 Training and Validation**
2. Chapter: Methods
    - 2.1 Hierarchical Clustering
    - 2.2 k-means
    - 2.3 DBSCAN
3. Chapter: Application
4. Chapter: Summary



Formal Definition - Clustering

- Elements $e \in E$
- Cluster c $\in C,\ with\ c\ \subseteq E\ and\ \bigcup_{c\,\in\,C} = E\ and\ \bigcap_{c\,\in\,C} = \emptyset$
- Representative $r_c = mean(c)$

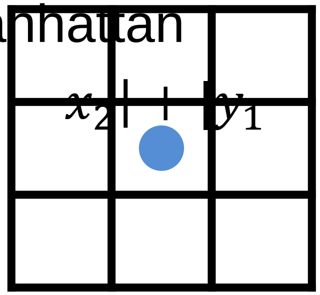
- $variability(c) = \sum_{e \in c} distance(r_c, e)^2$

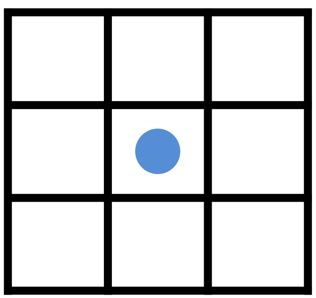
- Clustering: $Minimize \sum_{c \in C} variability(c)$

Formal Definition - Distance

Manhattan

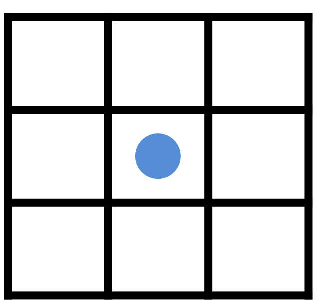
$|x_1 - x_2| + |y_1 - y_2|$





Euclidian

$\sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$



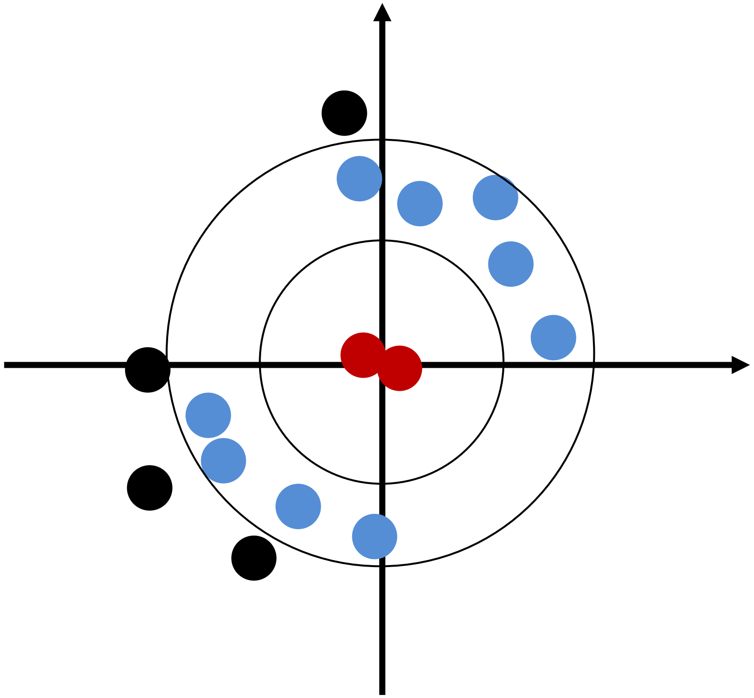
Chebyshev

$\max(|x_1 - x_2|, |y_1 - y_2|)$

Recap Classification

- Classification
  - Labeled training data (supervised)
  - Given classes

- Example: Dart
  - Shooting a target
  - 3 classes for points

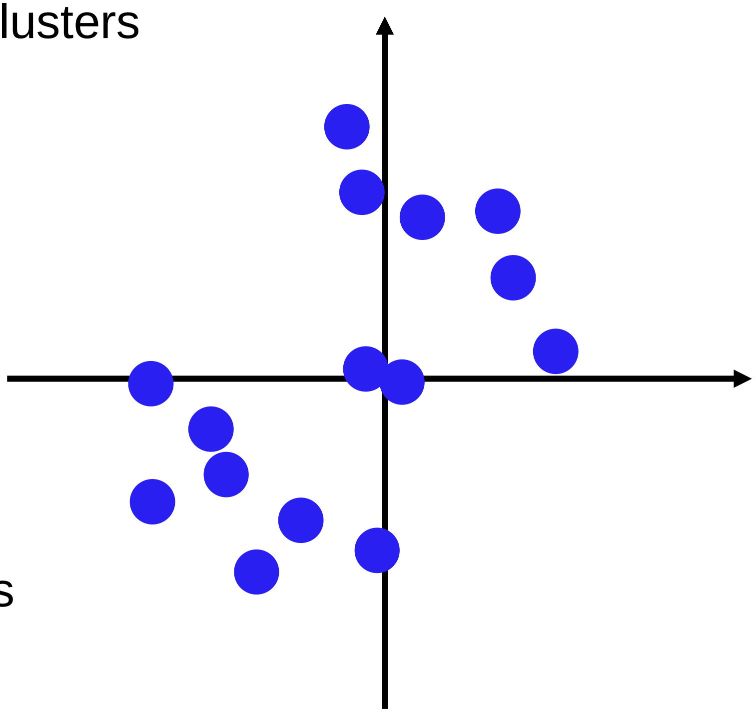


Clustering

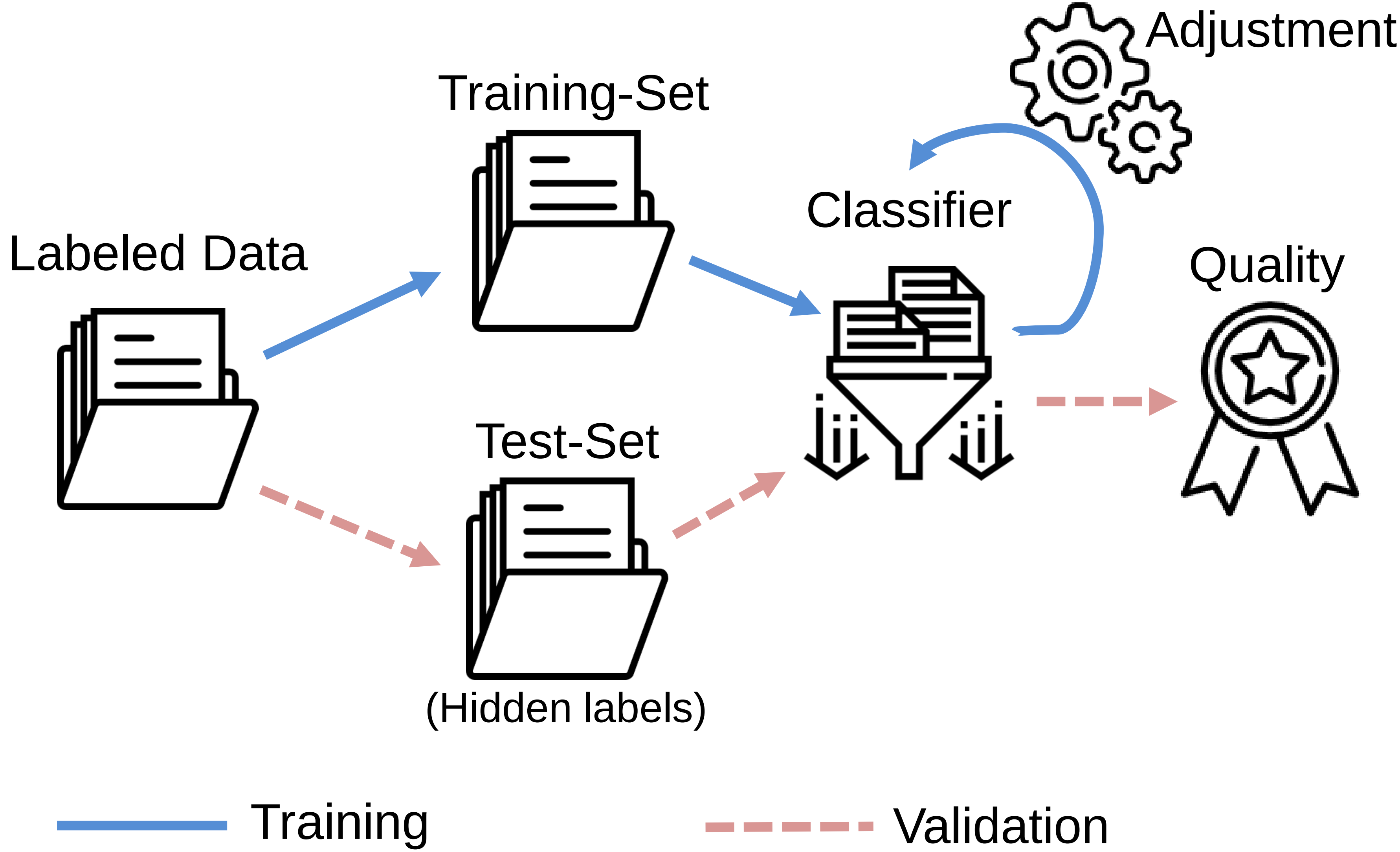
- Grouping a set of data objects into clusters
  - Cluster: a collection of elements
  - Similar to one another within the same cluster
  - Dissimilar to the objects in other clusters

- Difference to Classification
  - No given clusters/classes
  - Unsupervised learning

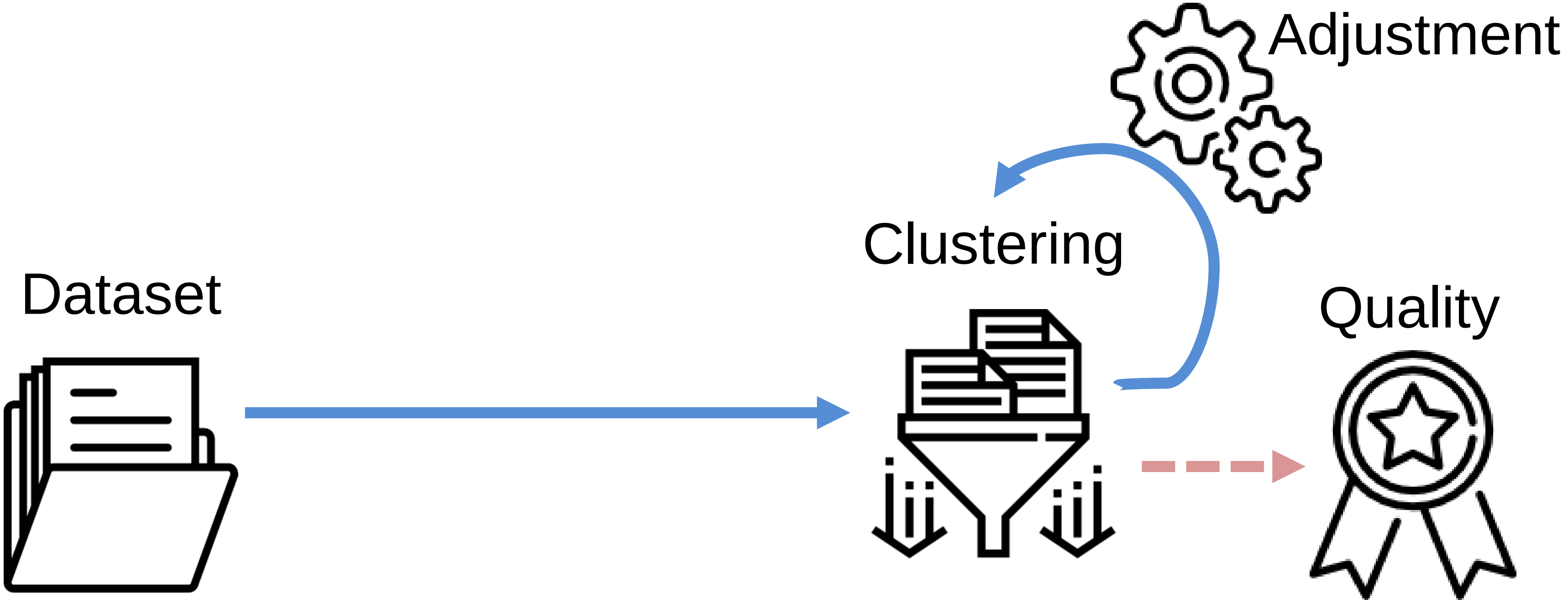
- Application
  - Get insights in large datasets
  - Preprocessing for other algorithms



Supervised Learning - Classification



Unsupervised Learning - Clustering

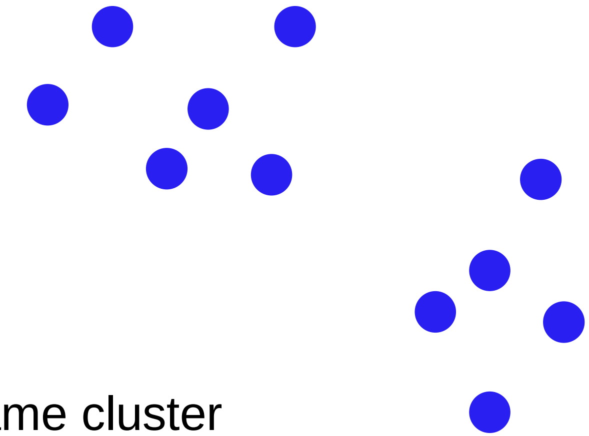


Training

Validation

Quality Measure of a Cluster

- Distances to representatives depend on k
  - $k = 2$: very large distances
  - $k = n - 1$: very small distances



- Similarity $sim(o)$
  - Within a cluster: $o \in a \in C$
  - Average distance to all elements within the same cluster
  - $sim(o) = \frac{1}{|a|} \sum_{e \in a} distance(o, e)$
- Dissimilarity $dsim(o)$
  - To other clusters: $e \notin b \in C$
  - Average distance to all elements of the second closest cluster
  - $dsim(o) = \min\limits_{c \neq a} (\frac{1}{|c|} \sum_{e \in c} distance(o, e))$

Quality Measure of a Cluster

- Silhouette coefficient
  - $s(o) = \frac{dsim(o) - sim(o)}{\max\{sim(o), dsim(o)\}}$
  - $if\ sim(o) = dsim(o) = 0, then\ s(o) = 0$
  - $s(o) \in [-1, 1]$
  - $silh(c) = \frac{1}{|c|} \sum_{o \in c} s(o)$
  - $silh(E) = \frac{1}{|E|} \sum_{o \in E} s(o)$

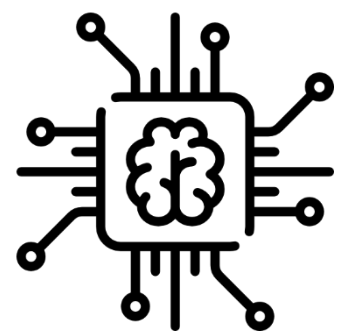
Supervised Learning: Classification

Johannes Betz / Prof. Dr. Markus Lienkamp / Prof. Dr. Boris Lohmann

(Jan Cedric Mertens, M.Sc.)

Agenda

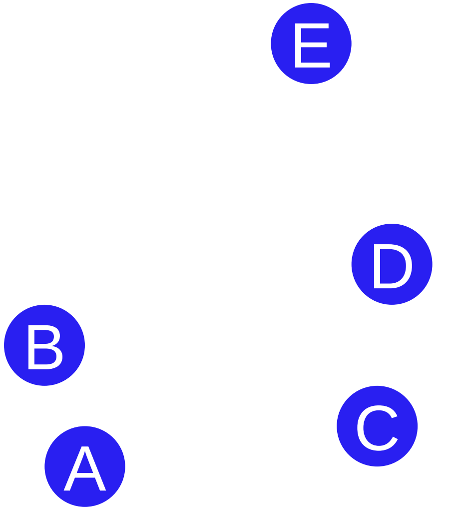
1. Chapter: Introduction
    - 1.1 Overview
    - 1.2 Training and Validation
2. Chapter: Methods
    - **2.1 Hierarchical Clustering**
    - 2.2 k-means
    - 2.3 DBSCAN
3. Chapter: Application
4. Chapter: Summary

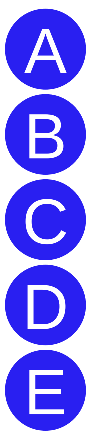


Hierarchical Clustering

1. Start with one cluster per element
2. Combine the two closest (most similar) clusters
3. Until all elements are in one cluster

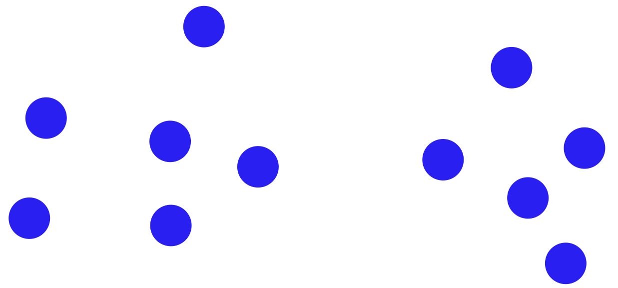
- Top down (divisive)/Bottom up (agglomerative)



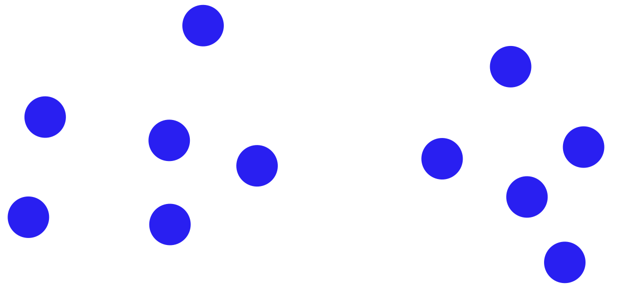


Distance between Clusters

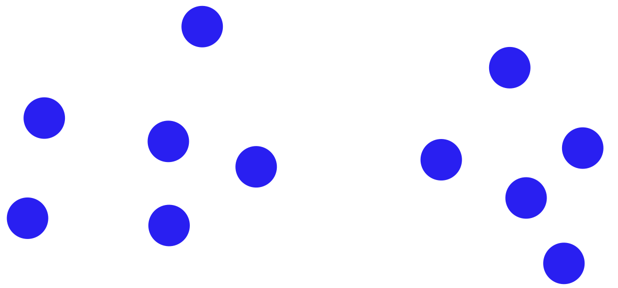
- Single Link
  - Smallest distance between two point of different clusters



- Complete Link
  - Largerst distance between two points of different clusters

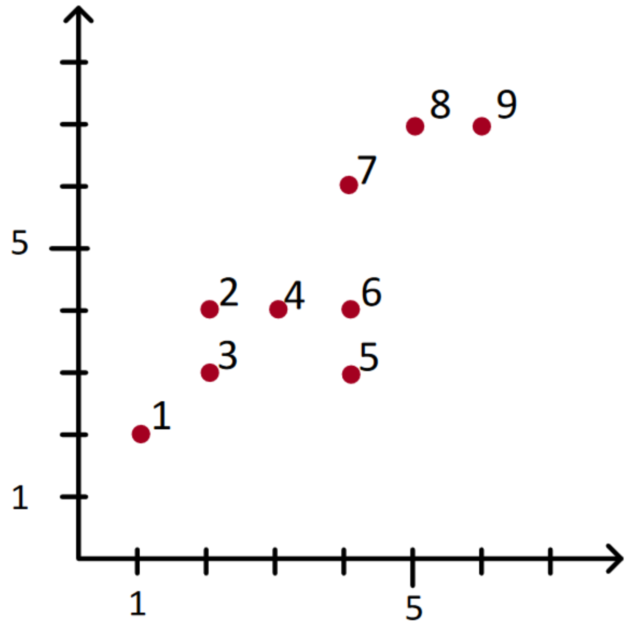


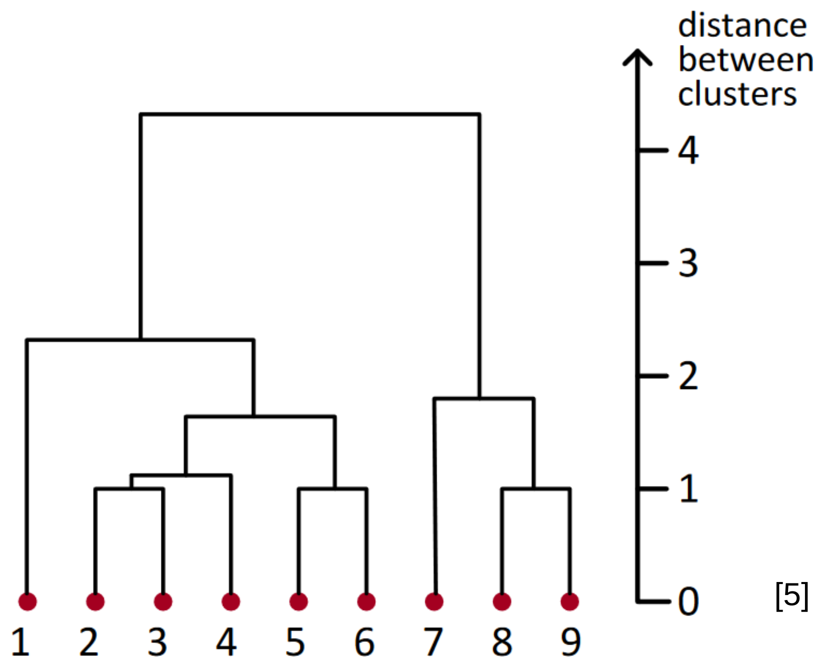
- Average Link
  - Average distance between all points of one cluster to all points of a different cluster



Dendrogram

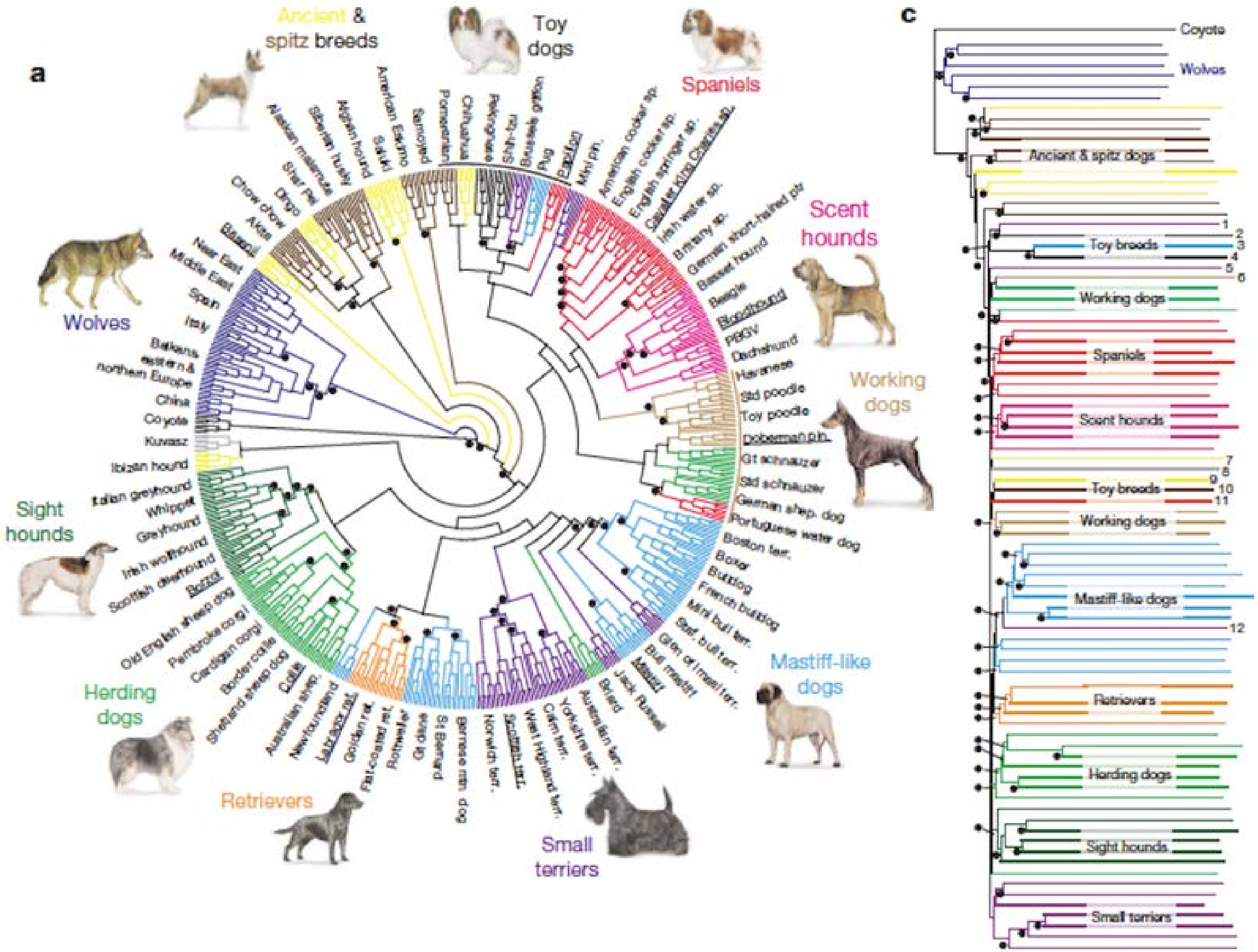
- Root: Cluster with all points
- Leaf: Cluster with one point
- Edges: Combine two clusters
- Depth: Distance between two combined clusters





[5]

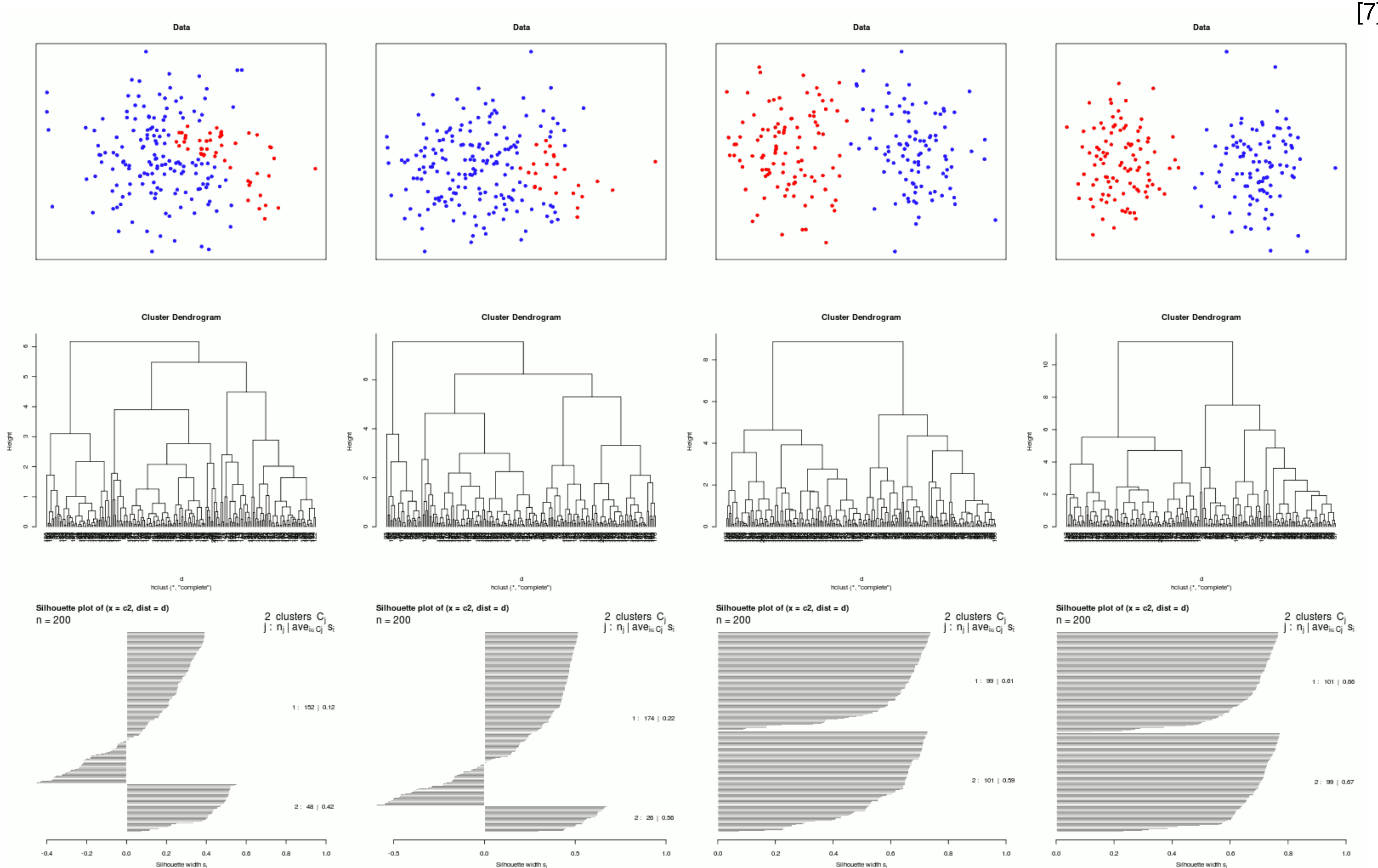
Dendogram - Example



[6]

Hierarchical Clustering - Silhouette coefficient

[7]



Hierarchical Clustering - Example

	BOS	NY	CHI	DEN	SF	SEA
BOS	0	206	963	1949	3095	2979
NY		0	802	1771	2934	2815
CHI			0	966	2142	2013
DEN				0	1235	1307
SF					0	808
SEA						0

[8]

Discussion Hierarchical Clustering

- Pro:
  - **Generic**: No cluster number or parameters must be defined
  - **Visualization**: E.g dendrogram shows hierarchy
  - **Hierarchy**: Relationship between clusters
  - **Deterministic:** Generates always the same clusters

- Contra:
  - **Scalability:** Runtime $\mathcal{O}(n^3)$
  - **Choice:** The final cluster must be selected from the hierarchy

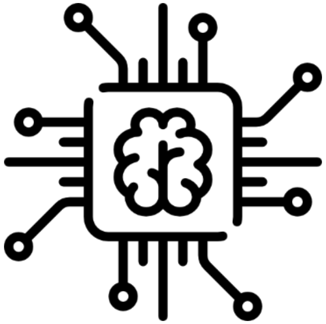
Supervised Learning: Classification

Johannes Betz / Prof. Dr. Markus Lienkamp / Prof. Dr. Boris Lohmann

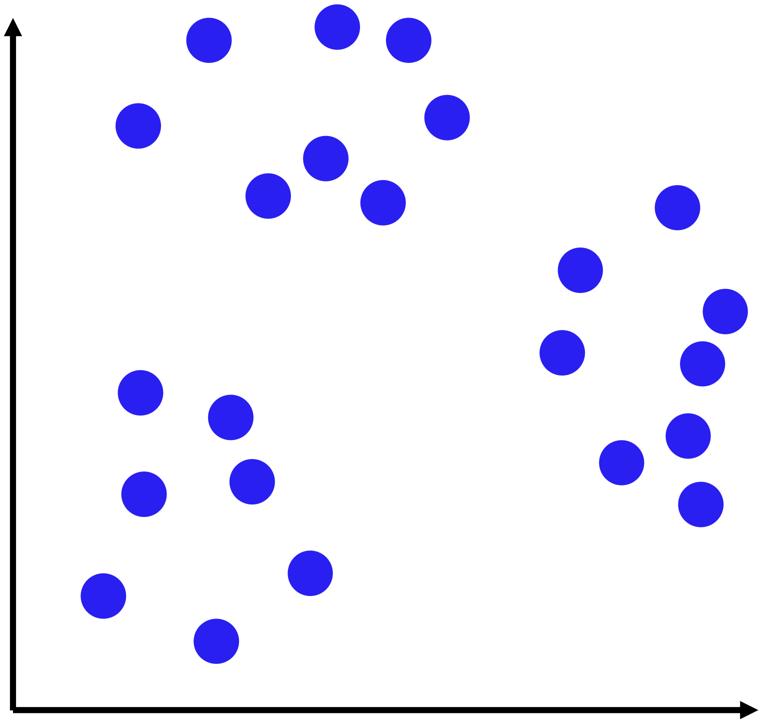
(Jan Cedric Mertens, M.Sc.)

Agenda

1. Chapter: Introduction
    1.1 Overview
    1.2 Training and Validation
2. Chapter: Methods
    2.1 Hierarchical Clustering
    2.2 k-means
    2.3 DBSCAN
3. Chapter: Application
4. Chapter: Summary



K-Means - Basic Idea



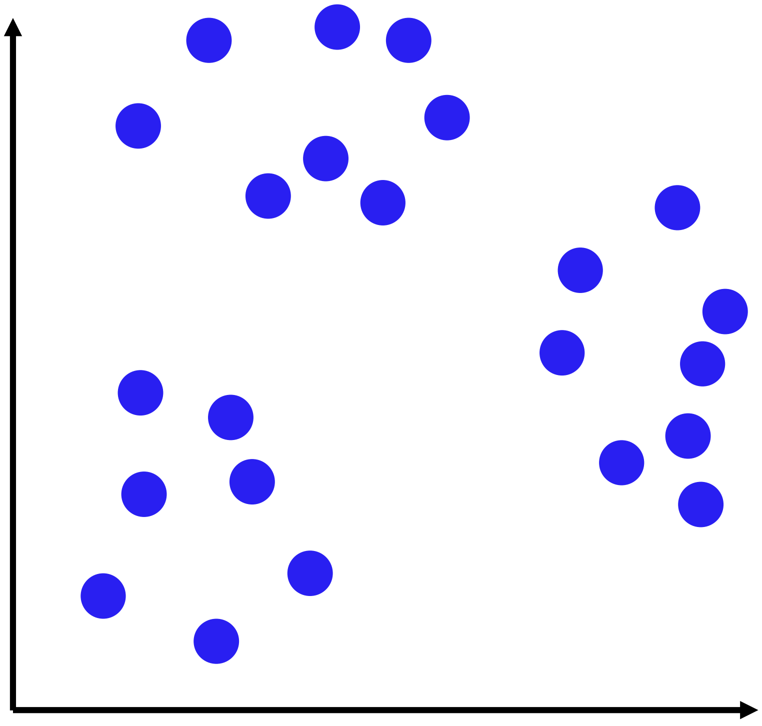
- Minimize squared distances to the cluster mean (variability)
- Minimize the summed variability of all clusters

  Large Sum → Poor clustering

  Minimal Sum → Optimal clustering

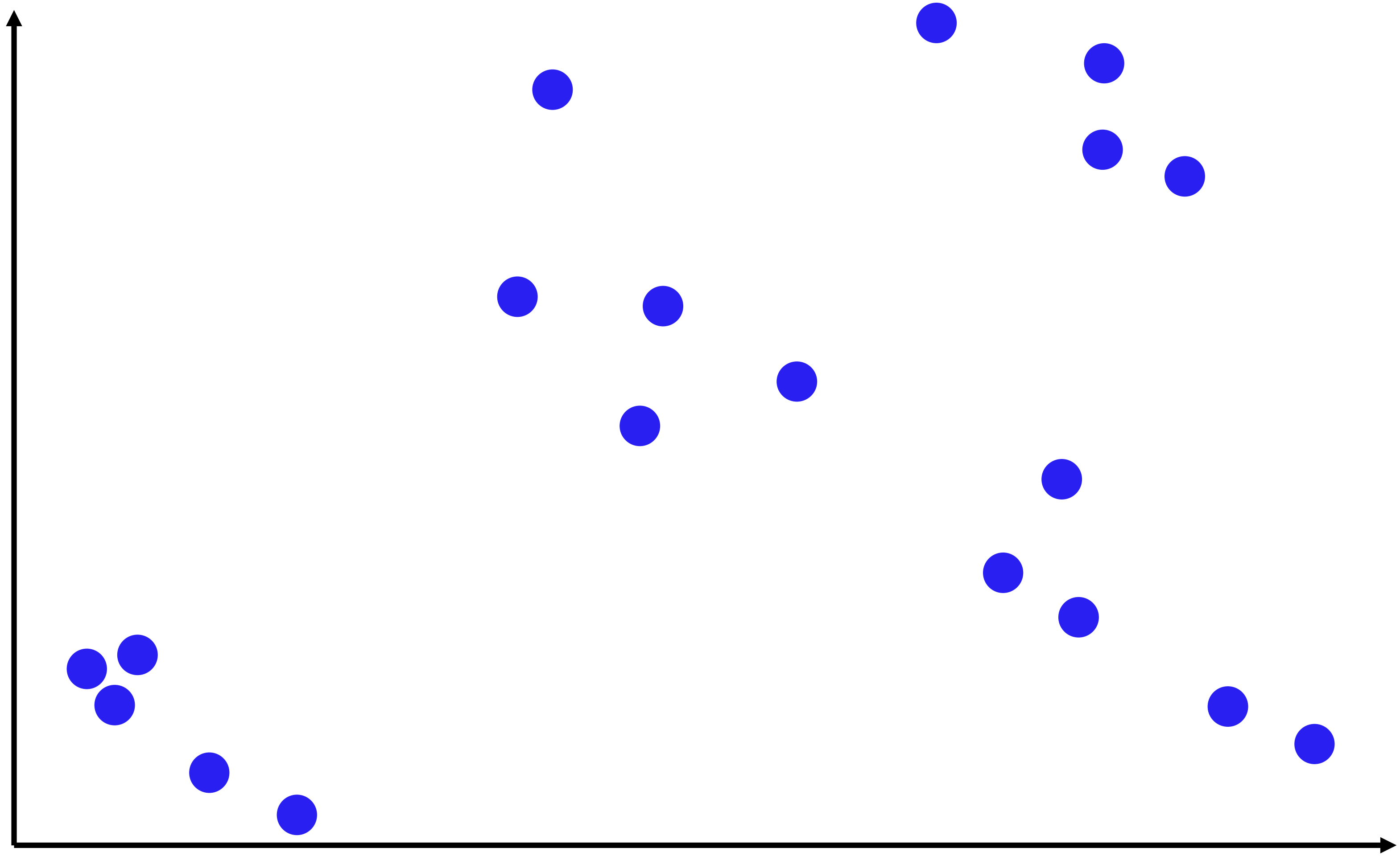
- Computationally challenging
  - NP-hard

K-Means Algorithm (Lloyd)



- Given:
  - Number of desired clusters k
  - Dataset
- Initialization:
  - Choose k arbitrary representatives
- Repeat until stable:
  - Assign objects to nearest representative
  - Compute center of each cluster as new representative

K-Means Algorithm (Lloyd)

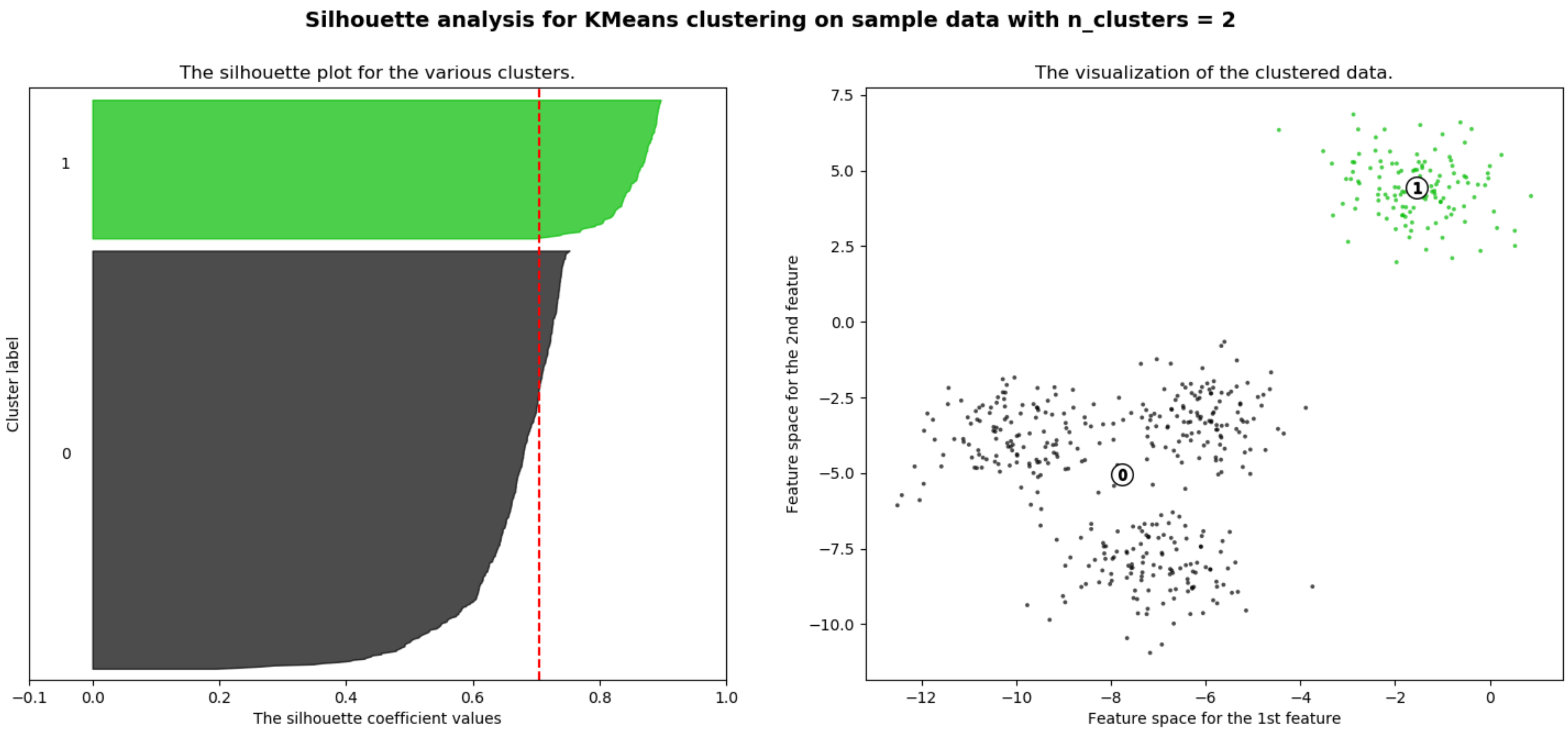


K-Means Algorithm – How to choose k?

- A priori knowledge of an expert
  - „There are five different types of bacteria“: k = 5

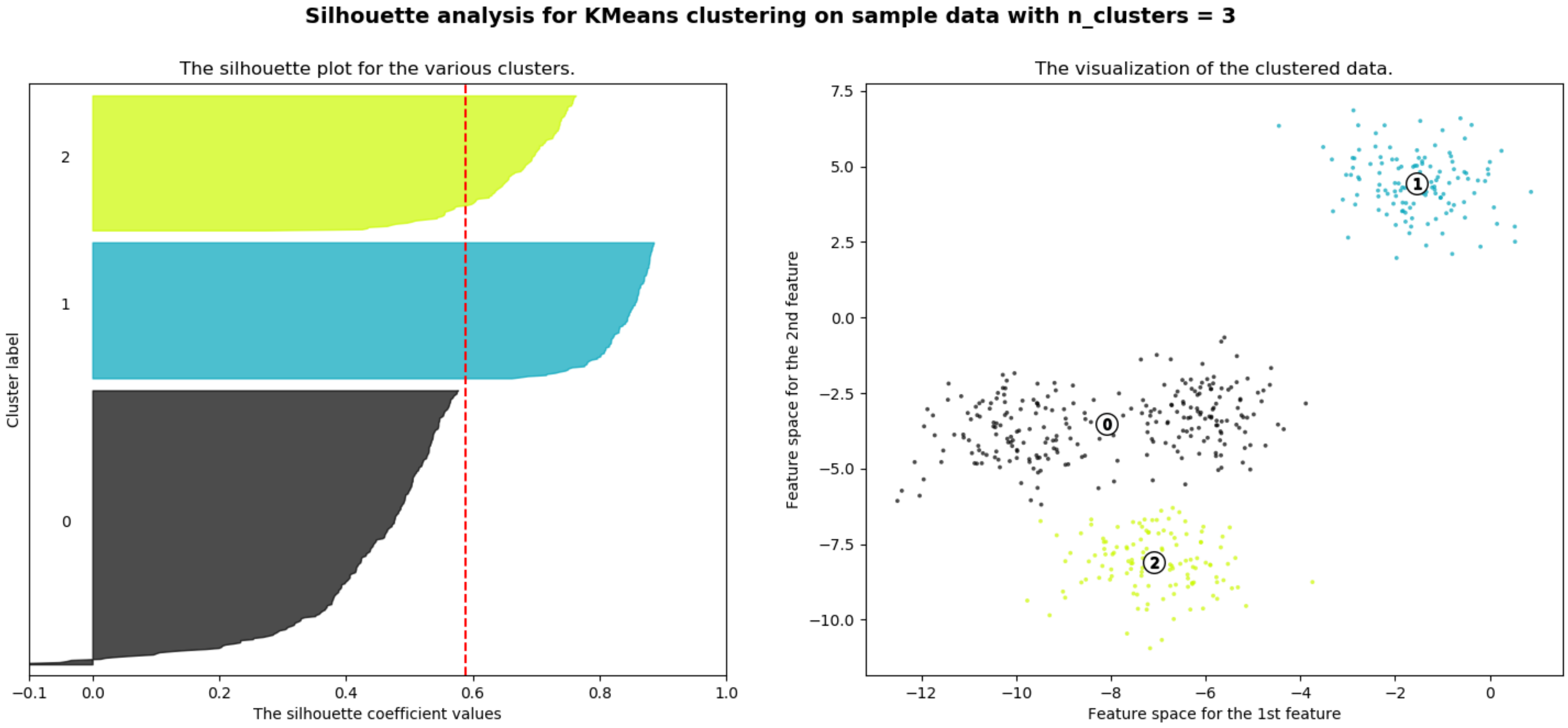
- Search for a good k
  - Naïve approach: Brute Force with k = 2 … n-1
  - Run hierarchical clustering on subset of data

K-Means Algorithm – How to choose k?



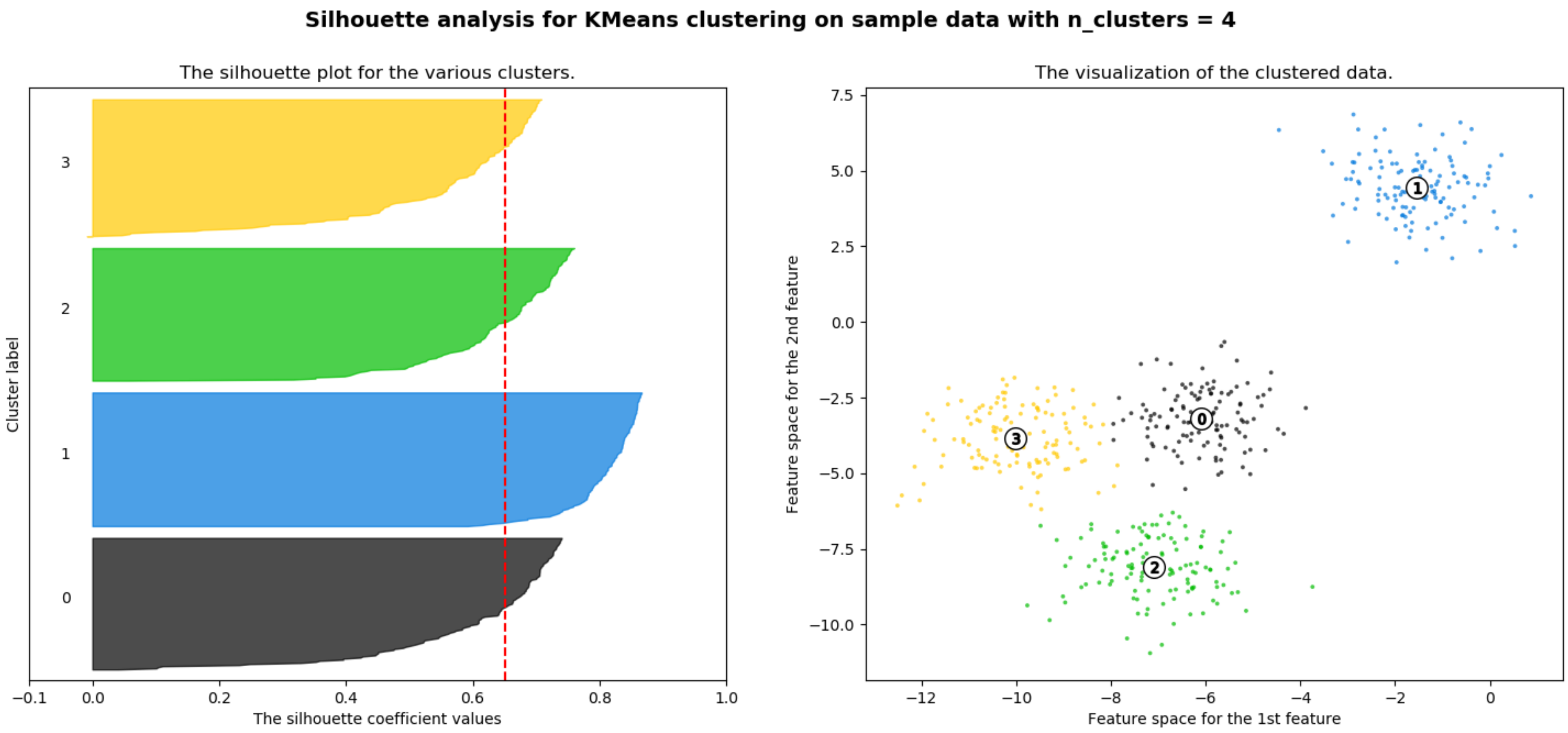
[9]

K-Means Algorithm – How to choose k?



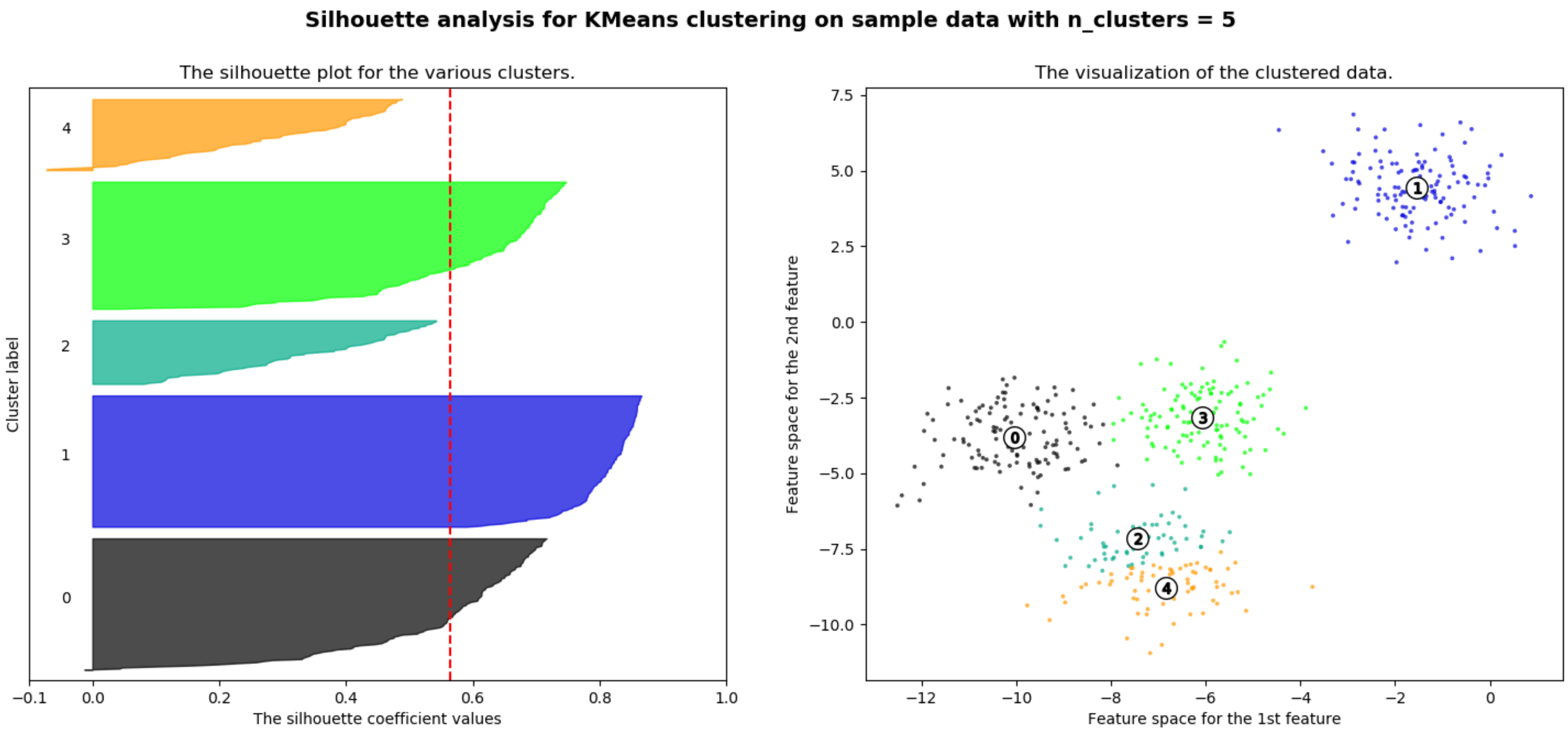
[9]

K-Means Algorithm – How to choose k?



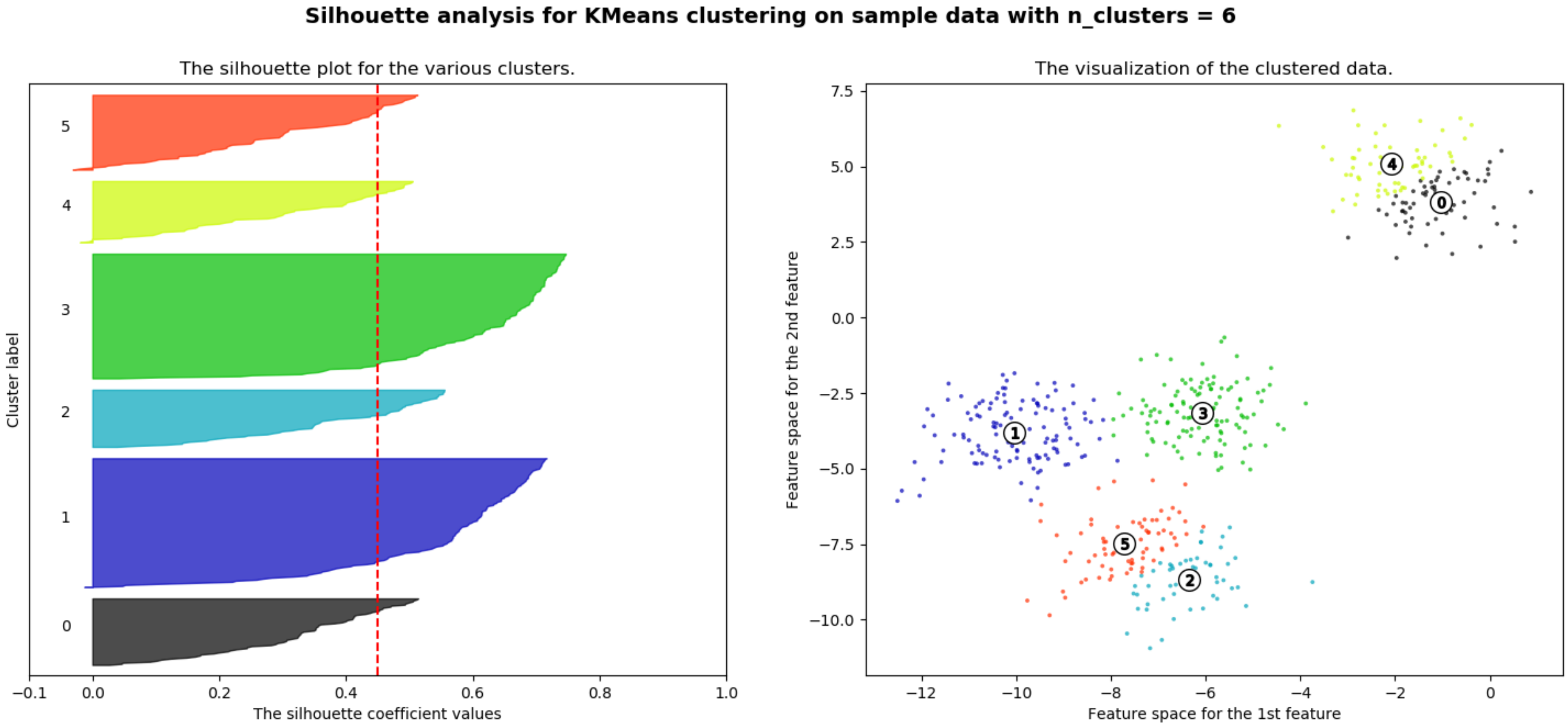
[9]

K-Means Algorithm – How to choose k?



[9]

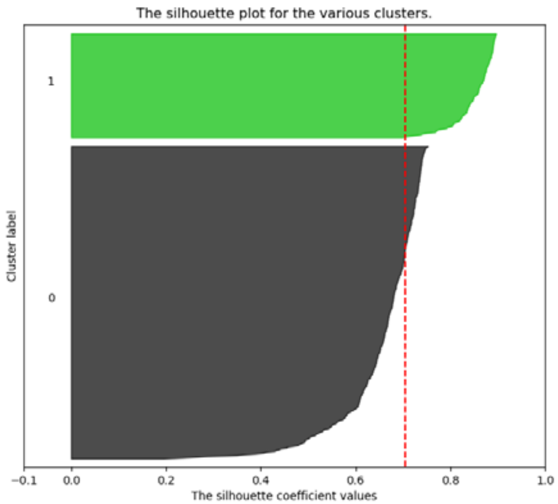
K-Means Algorithm – How to choose k?



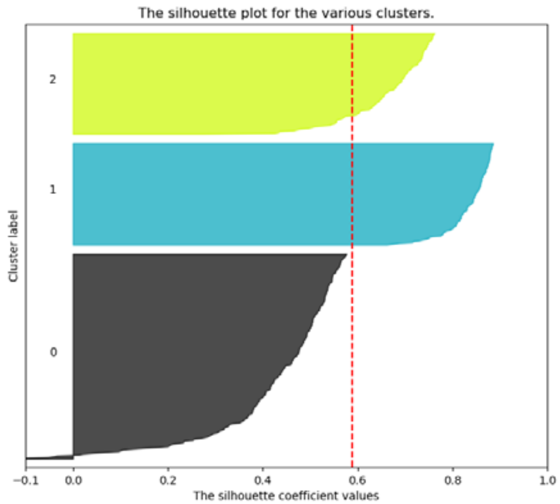
[9]

K-Means Algorithm – How to choose k?

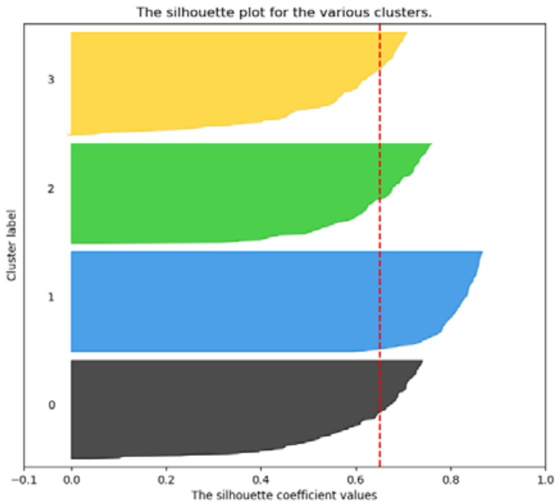
k = 2



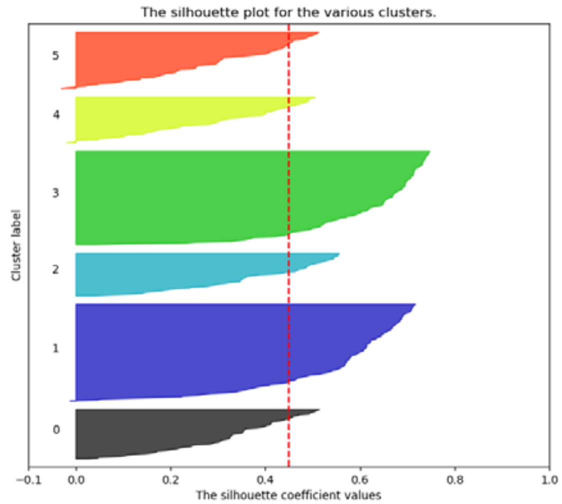
k = 3



k = 4

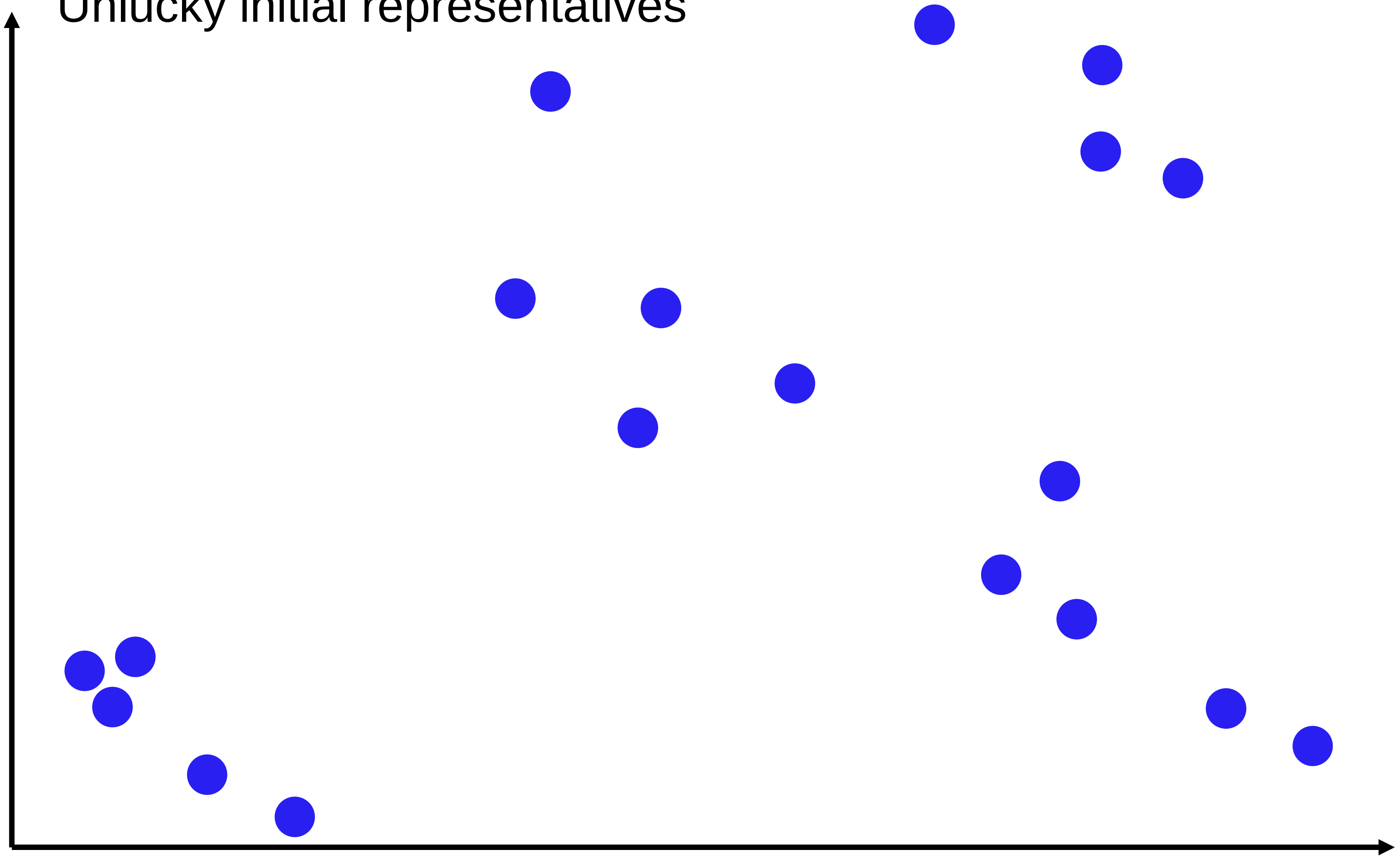


k = 5



K-Means Algorithm – How to handle randomness?

Unlucky initial representatives



K-Means Algorithm – How to handle randomness?

- Naïve approach
  - Get a small random subset D from E
  - Cluster D and use found representatives for initialization

- Improved approach
  - Get m small random subsets $A \dots M \subset E$
  - Cluster A to M and save representatives $R_A \dots R_M$
  - Cluster the merged set $AM = A \cup \cdots \cup M$, m times with $R_A \dots R_M$ as inititial representatives
  - Use the represenation ($R_A \dots R_M$) of the best clustering of AM as initial representation for E

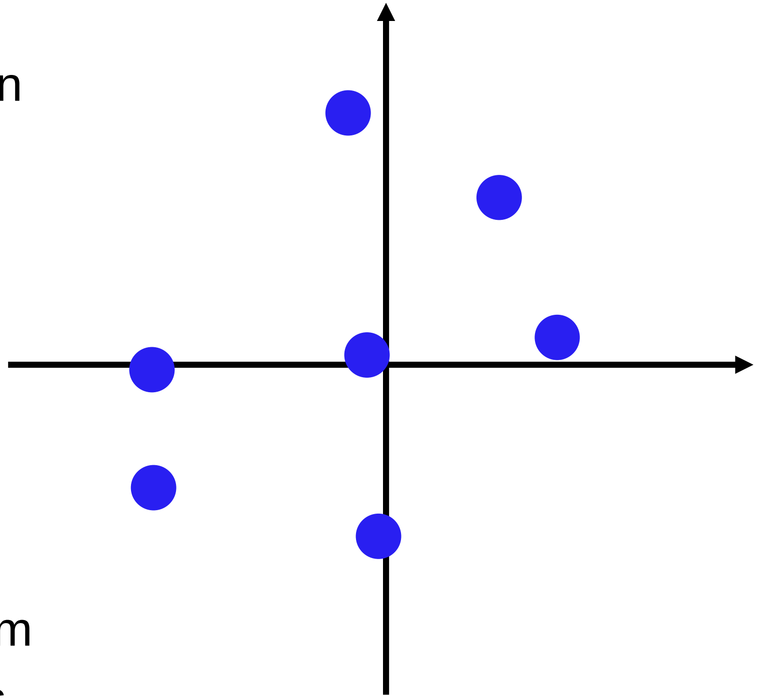
K-Means Example

[10]

Discussion K-Means

- Pro:
  - **Efficiency**: $\mathcal{O}(tkn)$ with typically k,t << n
    - n = #objects, k = #cluster, t = #iterations
  - **Implementation:** Easy to use

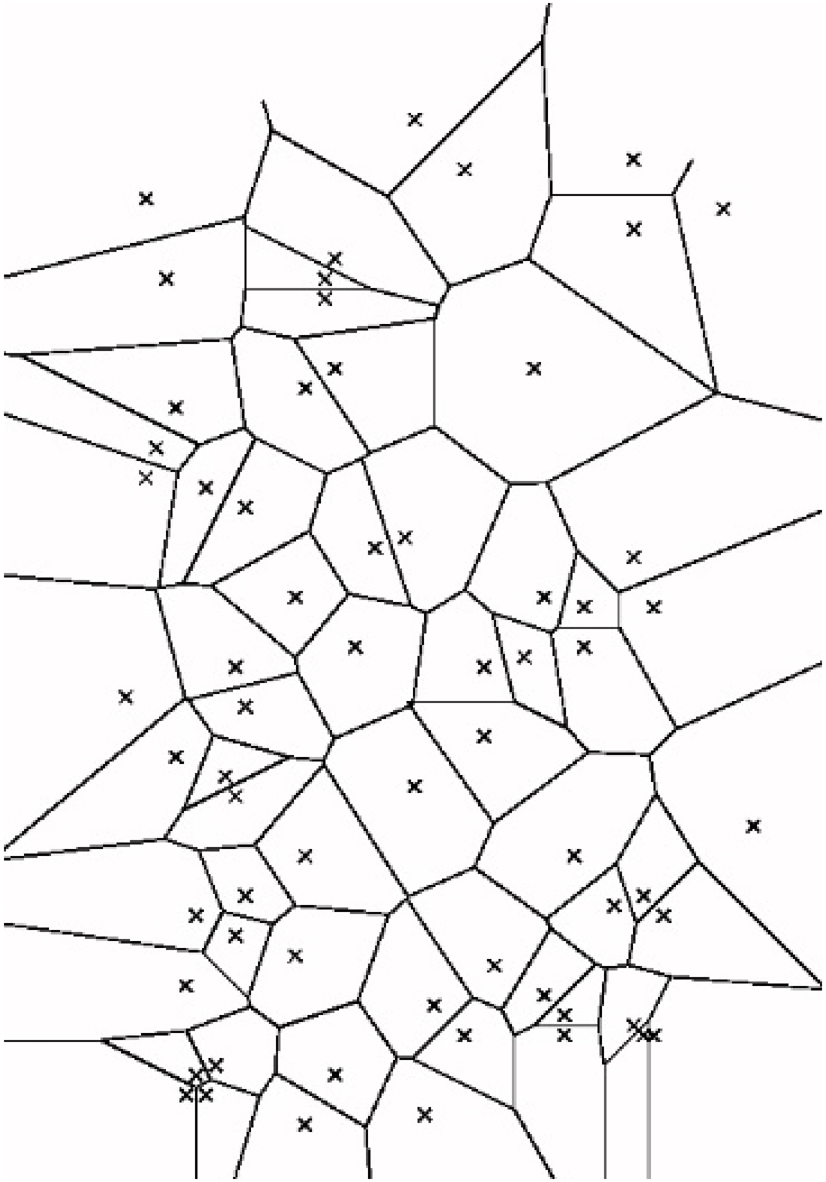
- Contra:
  - **Applicability**: mean must exist
  - **Noise:** Sensitive to outliers
  - **Specification:** k must be defined
  - **Initialization:** Might run in local optimum
  - **Cluster Form:** Convex space partitions



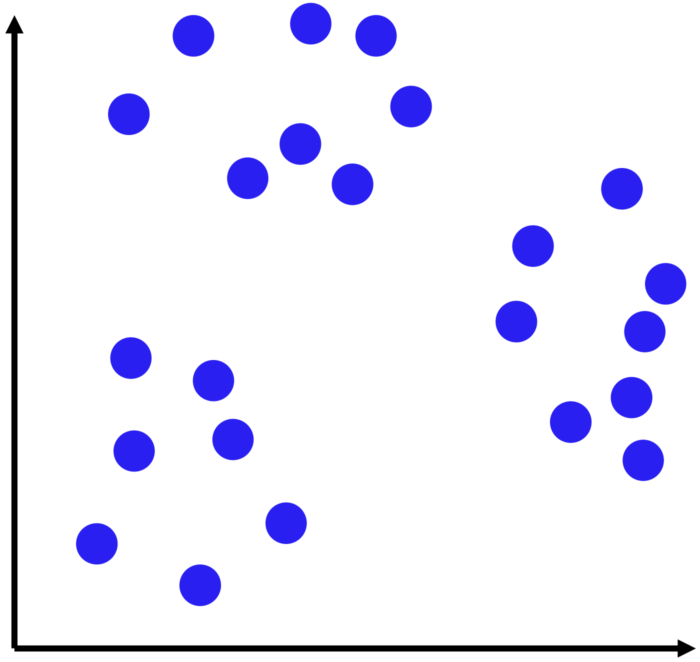
Voronoi Model

- The Voroni diagramm partiones the space in Voroni cells for each point p

- The Voroni cell for point p covers the area which nearest data point is p



Variants - K-Medoids, K-Median Clustering



- Representative: Mean → Object from cluster
  - Means do not always exist
- Distance: squared distance → normal distance
  - Influence of outliers is reduced
- Two variants for representative:
  - Medoid: Object in the middle
  - Median: Artificial object in the middle
- Basic idea:
  - Minimal distance between the objects of a cluster to its representative

Discussion k-Means, k-Medoid & k-Median

	K-means	K-medoid	K-median
data	Numerical data (mean)	metric	ordered attributed data
efficiency	High O(tkn)	Low O(tk(n-k)^2)	High O(tkn)
Senitivity to outlivers	High	Low	Low

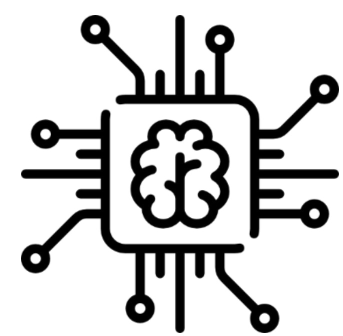
- Pro
  - **Implementation:** Easy to use
- Contra
  - **Specification:** k must be defined
  - **Cluster Form:** Convex space partitions
  - **Initialization:** Might run in local optimum

Supervised Learning: Classification
Johannes Betz / Prof. Dr. Markus Lienkamp / Prof. Dr. Boris Lohmann

(Jan Cedric Mertens, M.Sc.)

Agenda

1. Chapter: Introduction
   - 1.1 Overview
   - 1.2 Training and Validation
2. Chapter: Methods
   - 2.1 Hierarchical Clustering
   - 2.2 k-means
   - **2.3 DBSCAN**
3. Chapter: Application
4. Chapter: Summary

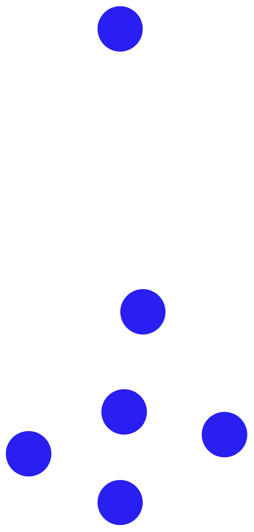


Density Based Clustering - DBSCAN

- Density-Based Spatial Clustering Application with Noise

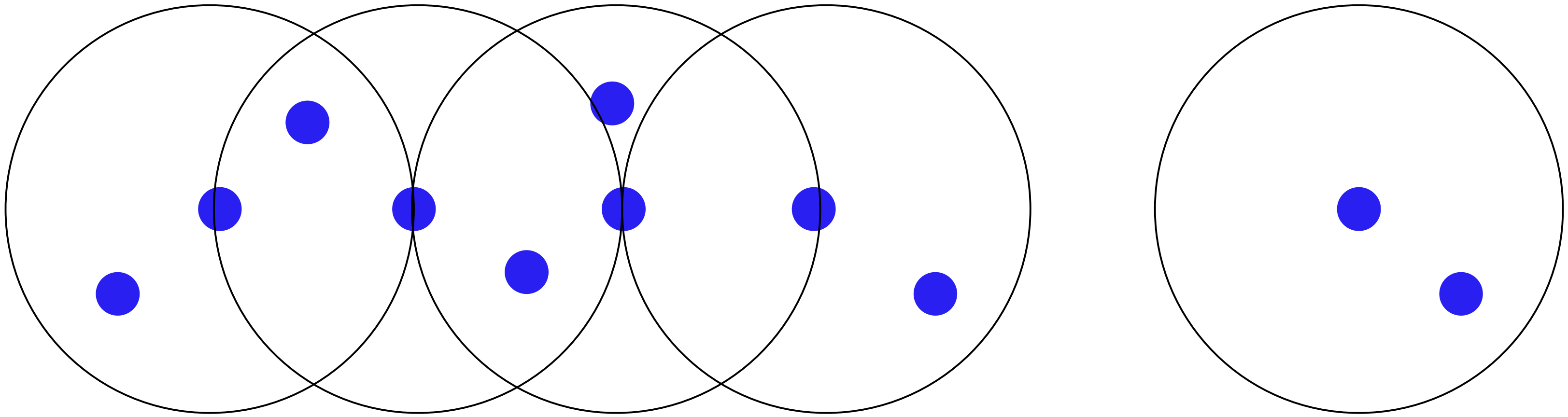
- Two parameters
  - $\mathcal{E}$-radius neighborhood
  - Minimum Points

- Three Point-classes
  - Core
  - Border
  - Outlier

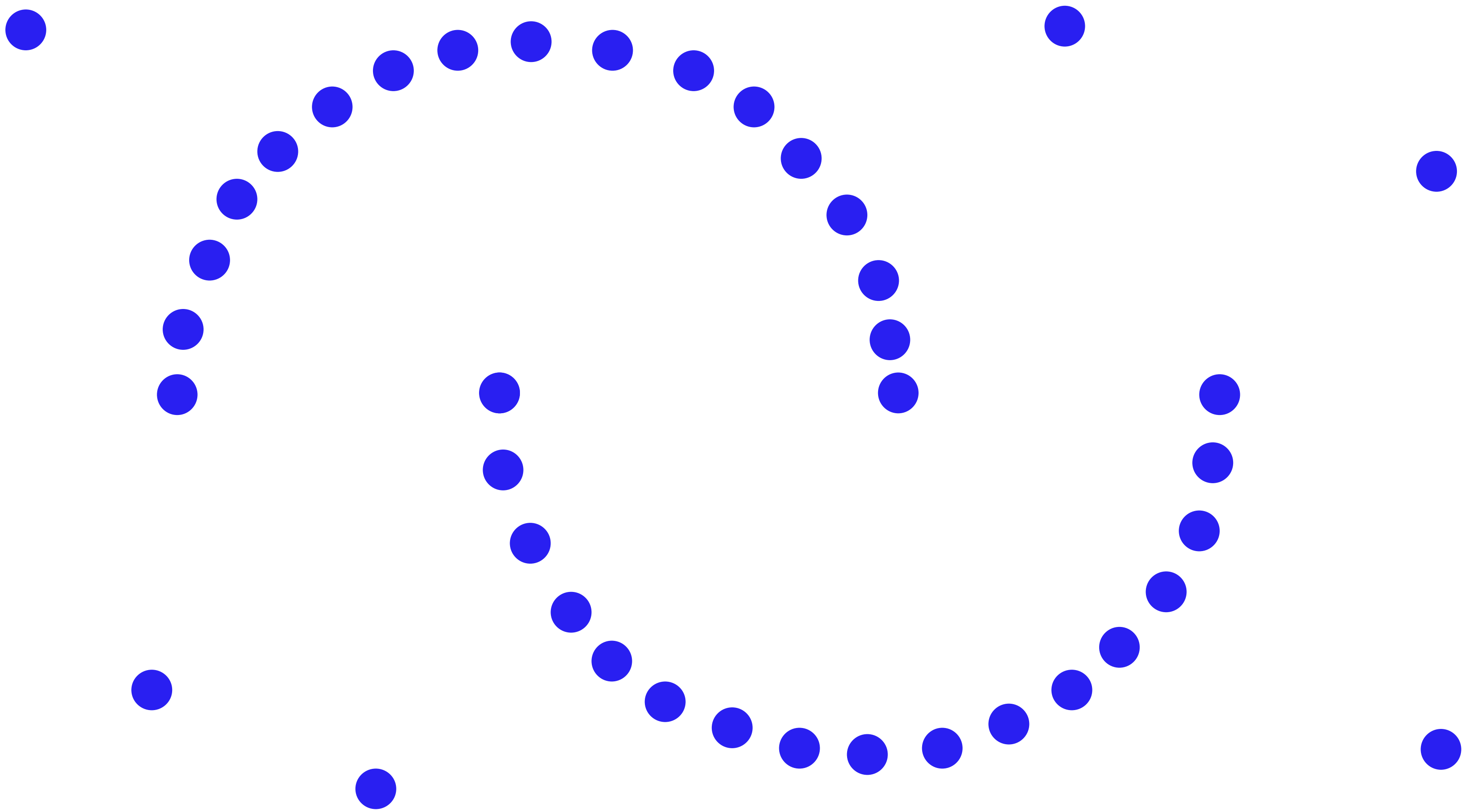


DBSCAN – Density Reachability

- p_n is „reachable“ from p_1, if there is a path p_1 ... p_n where each p_i on the path must be a core point, except for p_n



DBSCAN – Example



Discussion DBSCAN

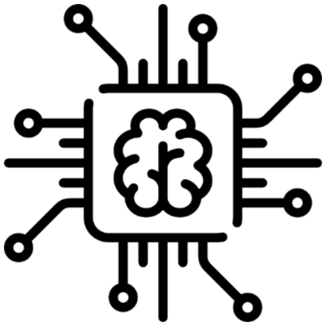
- Pro:
  - **Cluster Form:** Arbitrary space partitions
  - **Specification:** k is determined automatically
  - **Noise:** Separates clusters from noise
  - **Efficiency**: DBSCAN $\mathcal{O}(n^2)$

- Contra:
  - **Specification:** Parameters difficult to determine
  - **Sensitivity:** Very sensitive to parameter changes

Supervised Learning: Classification
Johannes Betz / Prof. Dr. Markus Lienkamp / Prof. Dr. Boris Lohmann

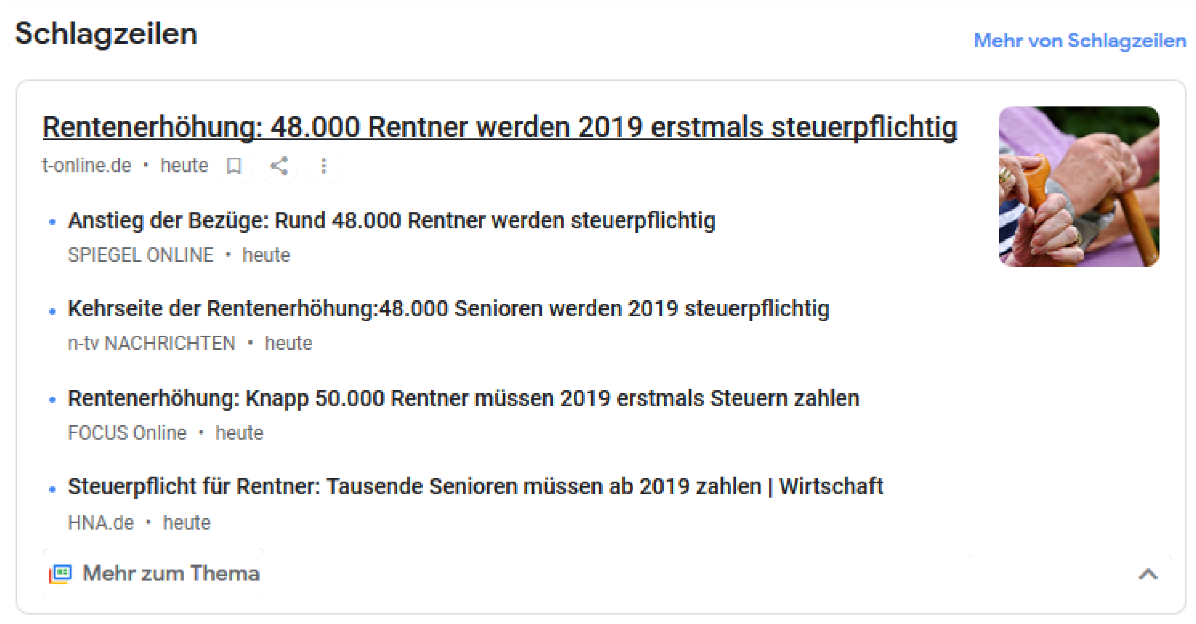
(Jan Cedric Mertens, M.Sc.)

Agenda



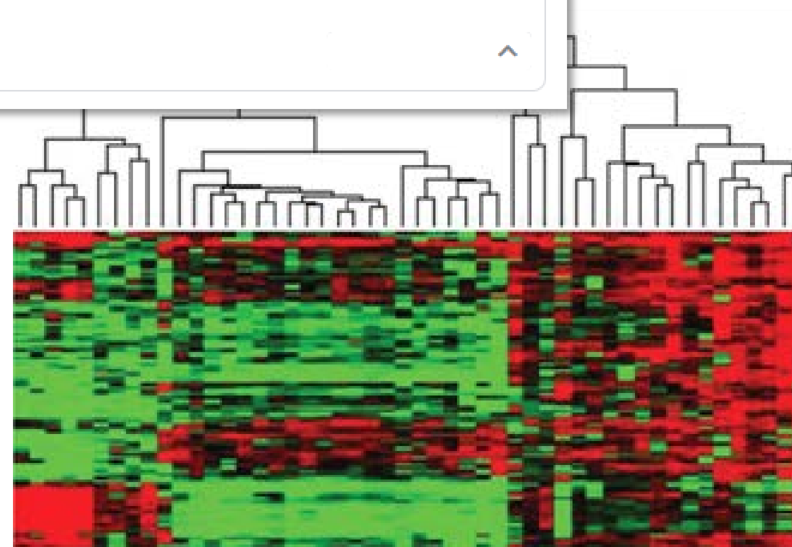
cluster_analysis.jpg

Applications



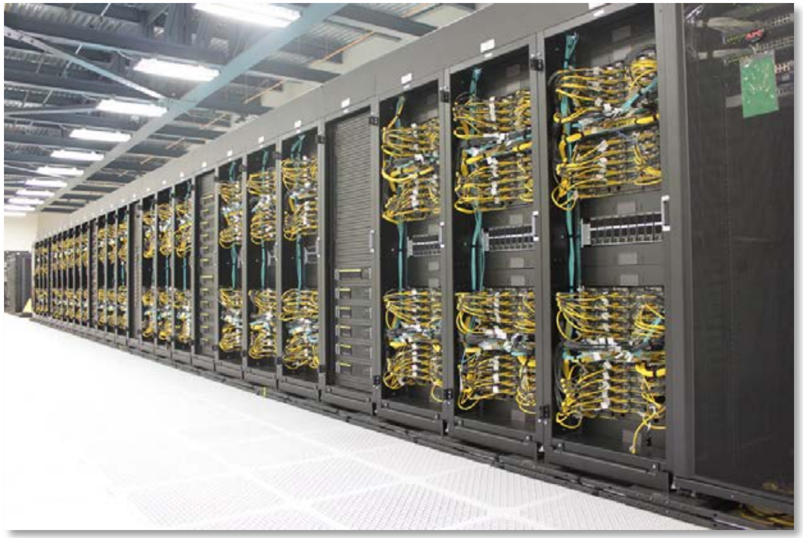
Google News

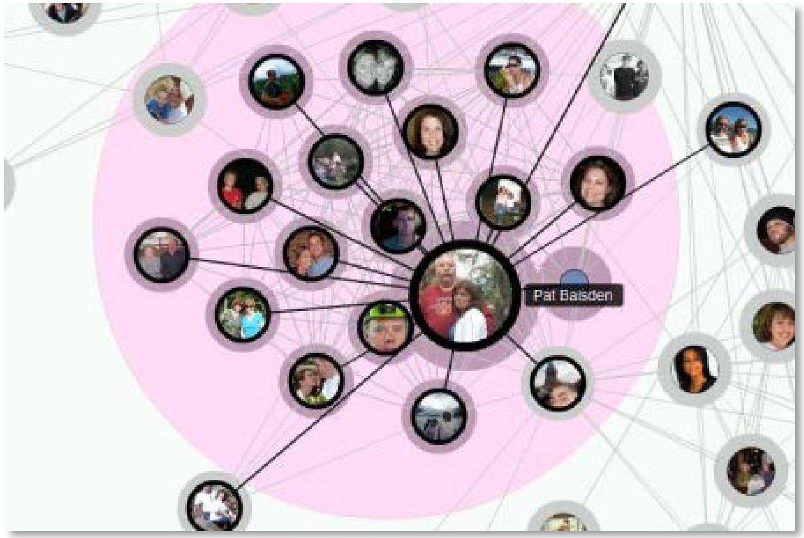
Genome Patterns



[14]

Applications

[11] 



[12]

Computing Cluster

Sozial Network

Market Segmentation

[13]

Application

- Customer Clustering
  - Amazon: Product suggestion (personalised advertisment)
  - Netflix: Movie suggestion

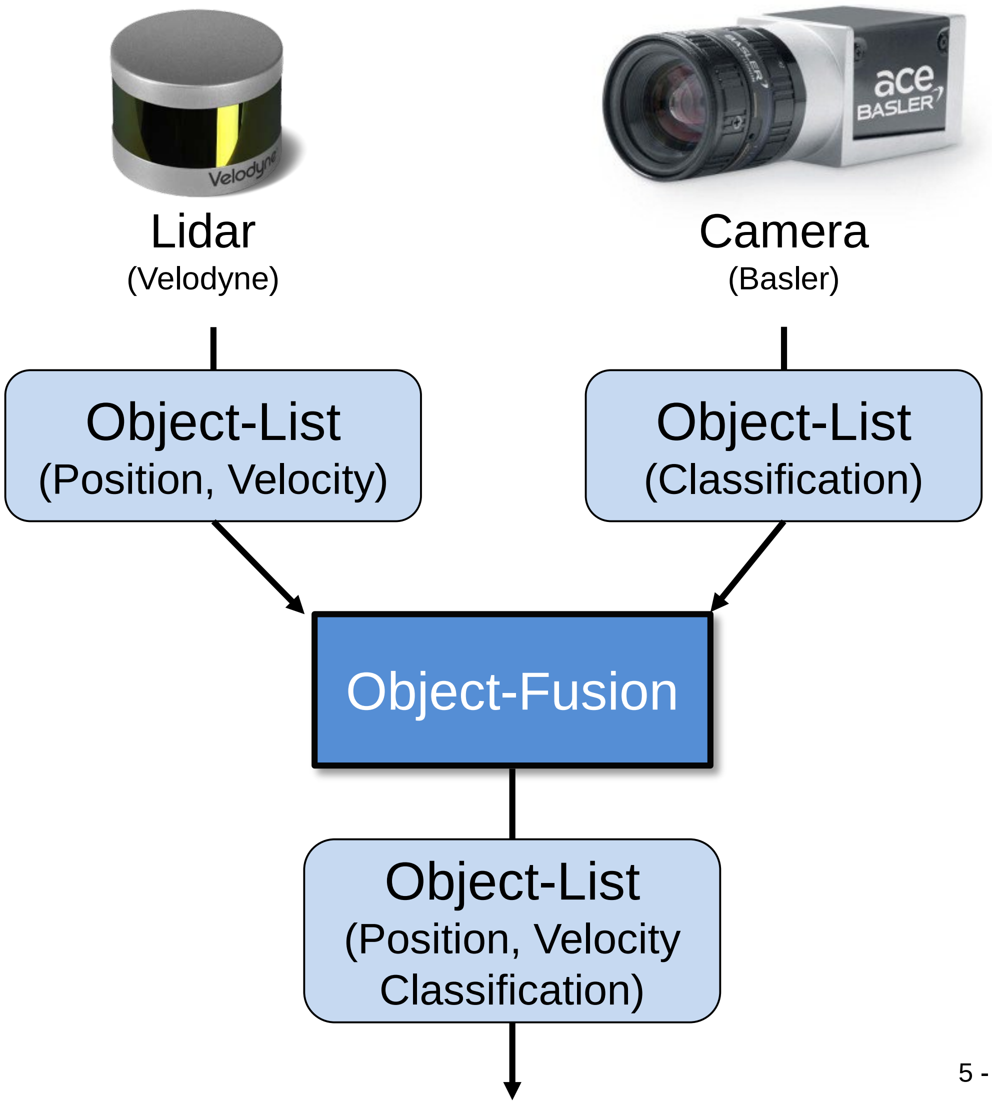
  - Netflix 1,000,000 $ challange from 2006

Clustering for automotive technology

- Traffic analysis
  - Collect mobility data of cars
    or density of certain regions
  - Use cluster algorithm to identify different groups
    - e.g. commuter, points of interest
  - Extract generalisation of trajectories and traffic flow
  - Use knowledge for city planing and to identify bottlenecks

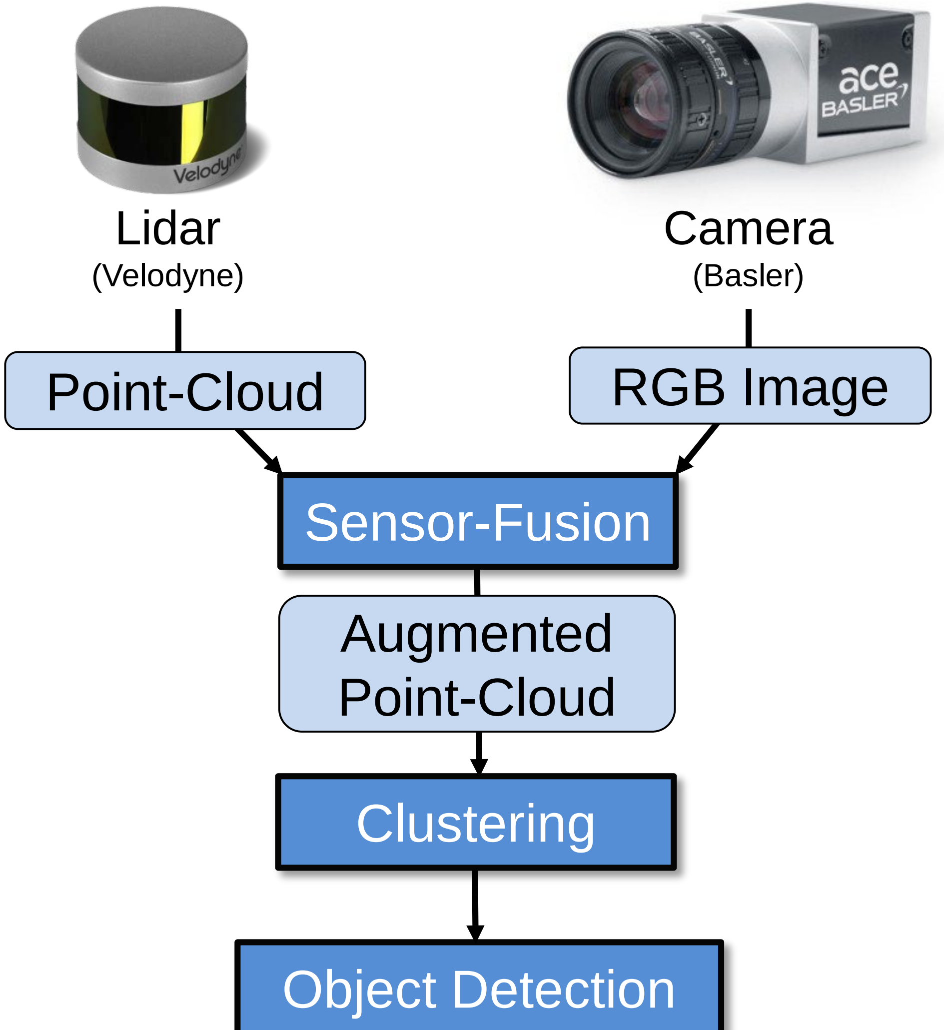
Clustering for automotive technology

- High Level Object-Fusion
  - Object-Detection based on limited data (only from one sensor)
  - Object-Fusion based on processed Object-List (already information loss)

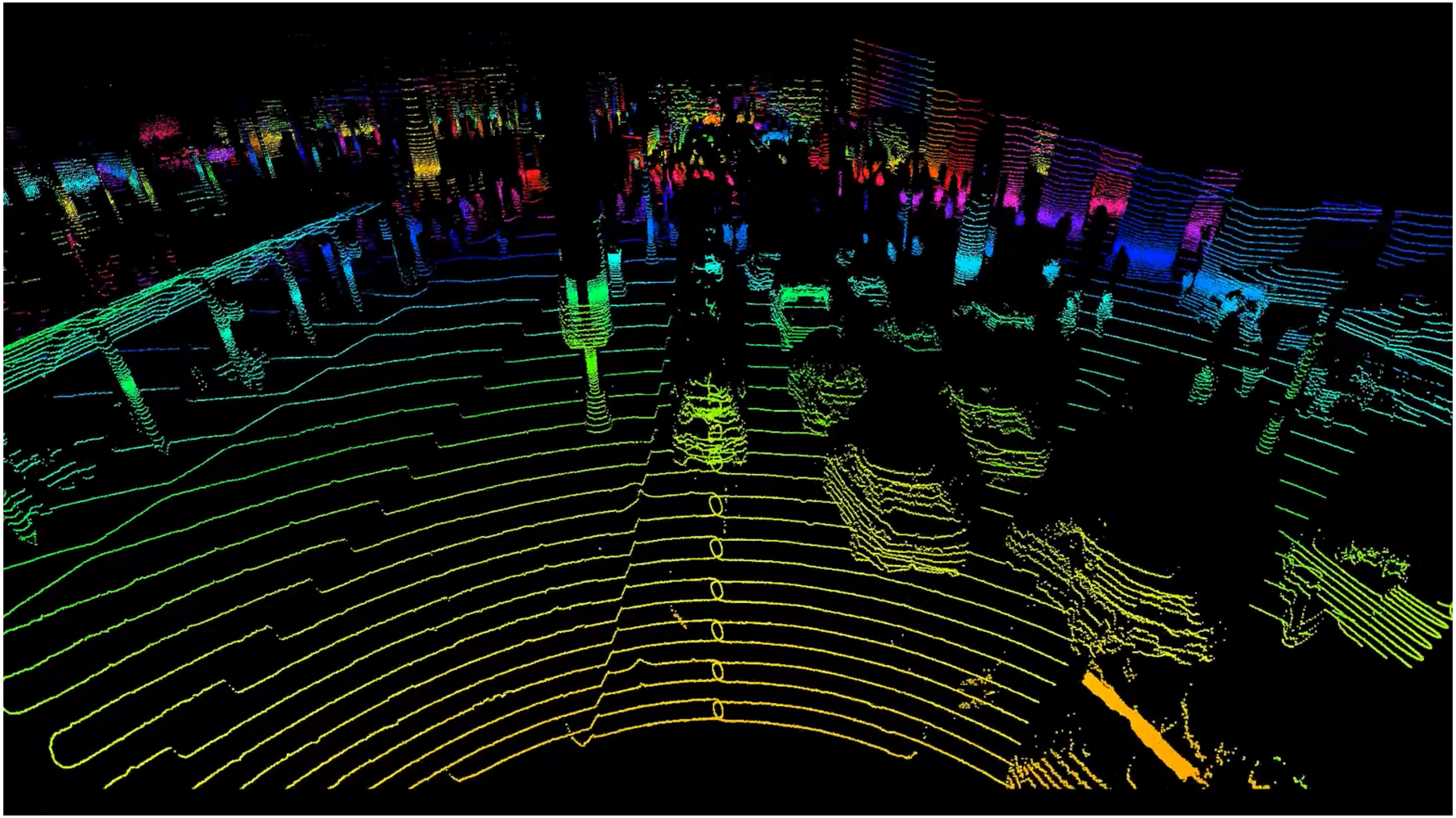


Clustering for automotive technology

- Low Level Object-Fusion
  - Overlay Lidar point-cloud with camera image
  - Find cluster in augmented pointcloud
  - Object Detection based on fused raw-data



Clustering for automotive technology



[15]

Clustering for automotive technology

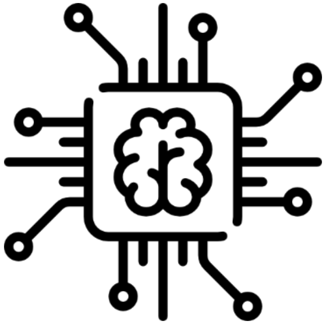
[4]

Supervised Learning: Classification

Johannes Betz / Prof. Dr. Markus Lienkamp / Prof. Dr. Boris Lohmann

(Jan Cedric Mertens, M.Sc.)

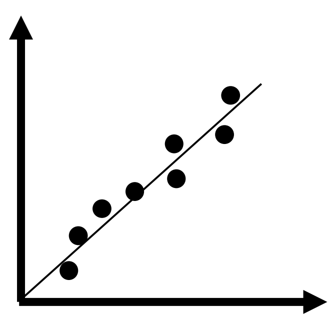
Agenda

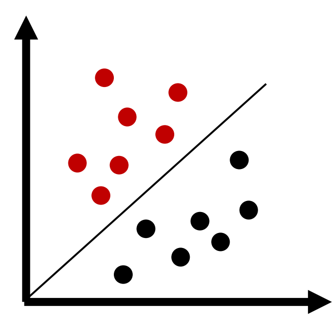


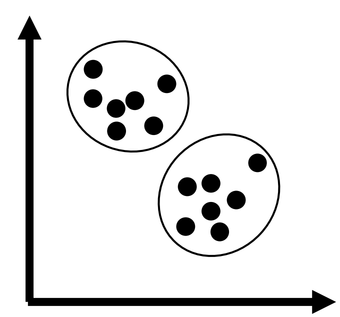
Summary

Pattern Recognition

Regression	**Classification**	Clustering
• Predict **continuous** valued output • Supervised	• Predict **discrete** valued output • Supervised	• Predict discrete valued output • **Unsupervised**







Summary

What did we learn today:

- **Clustering** is about finding groups in a dataset.
- Clustering is an **optimisation problem**.
- Elements **within** a cluster are **similar**.
- Elements form **different** clusters are **dissimilar**.
- The **distance** can be used to express similarity.
- Clustering is an **unsupervised** method, no labels are required.
- The **silhouette** can be used to express the **quality** of a cluster.
- Segmentation and Clustering are **interchangeable** terms.
- The concepts of hirarchical clustering, k-means and DBSCAN.
- Hirarchical clustering builds a dendogram.
- The number of desired clusters can be selected afterwards

Summary

What did we learn today:

- **K-means** is a fast but greedy and non determninistic algorithm.
- The **number of clusters** must be selected beforehand.
- Only **convex space** partitions can be generated.
- DBSCAN is a **density based** method and can deal with **noise**.
- Elements are classified as **core**, **border** or **outlier**.
- **Complex forms** can be grouped as clusters
- Clustering is applied as **preprocessing** or to find **coherences**.
- Wide range of **clustering applications**, but rarely as stand alone.
- Experts or classification methods **give clusters afterwards meaning**.

Sources

- [1] https://dailyillini.com/news/2017/09/28/students-reflect-race-affects-classroom-participation/
- [2] https://dictionary.cambridge.org/dictionary/english/cluster
- [3] https://www.deviantart.com/gttorres/art/Not-Another-High-School-Story-256629151
- [4] https://www.youtube.com/watch?v=xXWLXfMugkM
- [5] http://www.dbs.ifi.lmu.de/Lehre/KDD/WS1718/04_Clustering-3.pdf
- [6] http://www.instituteofcaninebiology.org/how-to-read-a-dendrogram.html
- [7] https://de.wikipedia.org/wiki/Silhouettenkoeffizient
- [8] https://ocw.mit.edu/courses/electrical-engineering-and-computer-science/6-0002-introduction-to-computational-thinking-and-data-science-fall-2016/lecture-slides-and-files/MIT6_0002F16_lec12.pdf
- [9] http://docs.w3cub.com/scikit_learn/auto_examples/cluster/plot_kmeans_silhouette_analysis/
- [10] https://www.youtube.com/watch?v=BVFG7fd1H30
- [11] https://computing.llnl.gov/tutorials/linux_clusters/
- [12] http://www.messersmith.name/wordpress/2009/10/12/visualizing-your-facebook-network-of-friends/
- [13] http://www.businessstudynotes.com/marketing/principle-of-marketing/discuss-market-segmentation-and-market-targetting/
- [14] http://www.discoveryandinnovation.com/BIOL202/notes/images/cluster_analysis.jpg
- [15] https://www.nytimes.com/2017/05/25/automobiles/wheels/lidar-self-driving-cars.html

Acknowledgment

- **Machine Learning (Stanford/Coursera)**
  - Andrew Ng

  [**https://**www.coursera.org/learn/machine-learning](https://www.coursera.org/learn/machine-learning)

- **Knowledge Discovery in Databases I (LMU)**
  - Prof. Dr. Peer Kröger

  http://www.dbs.ifi.lmu.de/cms/studium_lehre/lehre_master/kdd1718/index.html

- **Introduction to Computational Thinking and Data Science (MIT)**
  - Prof. Eric Grimson

  https://ocw.mit.edu/courses/electrical-engineering-and-computer-science/6-0002-introduction-to-computational-thinking-and-data-science-fall-2016